

UNIVERSIDADE FEDERAL DE SANTA CATARINA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO
CURSO DE DOUTORADO

GRÁFICOS DE CONTROLE PARA VARIÁVEIS
NÃO-CONFORMES AUTOCORRELACIONADAS

Suzana Leitão Russo

Tese submetida à Universidade Federal de Santa Catarina para obtenção do título de
Doutor em Engenharia de Produção.

FLORIANÓPOLIS
2002

Suzana Leitão Russo

GRÁFICOS DE CONTROLE PARA VARIÁVEIS NÃO-CONFORMES AUTOCORRELACIONADAS

Esta Tese foi julgada adequada para obtenção do Título de "Doutor em Engenharia", Especialidade em Engenharia de Produção e aprovada em sua forma final pelo Programa de Pós-Graduação em Engenharia de Produção.

Prof. Edson Pacheco Paladini
Coordenador do Curso de Pós-graduação

Banca Examinadora: _____

Prof. Robert Wayne Samohyl, Ph.D.
Orientador

Prof. João Neiva de Figueiredo, Dr.
Moderador

Prof^a. Maria Emília Camargo, Dra.

Prof. Paulo José Ogliari, Dr.

Prof. Fernando Menezes Campello de Souza, Ph.D.

RESUMO

Palavras-chaves: Autocorrelação, Gráficos de Controle, Variáveis Não-Conformes

Nesta tese restringe-se a analisar os gráficos de controle para variáveis contínuas e discretas. Os gráficos de controle convencionais de Shewhart, foram utilizados acrescidos de outros modelos adequados a transformações de observações autocorrelacionadas em observações que sejam independentes e normalmente distribuídas. Os dados utilizados para análise foram coletados na Indústria Têxtil Oeste Ltda., no município de Mondai – SC. As séries analisadas foram o número de ocorrências por tipo de não-conformidade em bobinas no setor de tecelagem (variável discreta e independente) e os valores da gramatura (denier) da fita de ráfia (variável contínua e explicativa) no setor de produção de fitas de polipropileno. Além de se verificar a autocorrelação dos dados, pode-se modelar as variáveis discretas através de modelos de regressão de Poisson e as variáveis contínuas através dos modelos Box e Jenkins e, com os resíduos obtidos utilizar os modelos de função de transferência para se identificar a existência de causalidade. A técnica proposta, de primeiro retirar a autocorrelação dos dados para depois ajustá-los mostrou-se satisfatória estatisticamente. Ao se estudar a autocorrelação dos dados gerou uma nova perspectiva de aprendizagem sobre o processo produtivo através das informações contidas na estrutura de autocorrelação, dos modelos Box e Jenkins e dos modelos de regressão de Poisson, aos quais eram ignorados pelo modelo clássico de monitoramento. A função de transferência empregada posteriormente nos resíduos obtidos permitiu a confirmação da causalidade da gramatura da fita com relação ao tecido produzido. Com isso houve um crescimento de informações para a correta tomada de decisão e pode-se detectar que houve uma melhora nos pontos de saída de controle.

ABSTRACT

Key-words: Autocorrelation, Control Charts, Nonconforming Items

In this thesis is restricted to analyze it the control charts for continuous and discrete variables. The conventional control charts of Shewhart, had been used increased of other adequate models the transformations of autocorrelation data in observation that are independent and are normally distributed. The data used for analysis had been collected in the Industry Têxtil Oeste Ltda. in the city of Mondai - SC. The analyzed series had been the number of occurrences for type of nonconformities in bobbins in the sector of weaving (discrete and independent variable) and the values of the weight (denier) of the ribbon of polypropylene (continuum and explanatory variable) in the sector of production of polypropylene ribbons. Beyond verifying the autocorrelation of the data, can adjust the discrete variable through models of regression of Poisson and the continuous variable through the models Box and Jenkins, and with the residues obtain we use the models of transference function to identify the causality existence. The technique proposal, of first removing the autocorrelation of the data for later adjusting revealed statistics satisfactory. To studying the autocorrelation of the data it generated a new perspective of learning on the productive process through the information contained in the structure of autocorrelation, the models Box and Jenkins and the models of regression of Poisson, which were ignored by the classic model of adjustment. The function of transference employed later in the gotten residues allowed the confirmation of the causality of the weight of ribbon of polypropylene with relation to the fabric produced. With this it had a growth of information for the correct taking of decision and can be detected that it had an improvement in the leaving points of control.

SUMÁRIO

RESUMO	iii
ABSTRACT	iv
LISTA DE FIGURAS	viii
LISTA DE TABELAS	xi
LISTA DE SÍMBOLOS E ABREVIATURAS	xiii
1. INTRODUÇÃO	1
1.1 TEMA E PROBLEMA	1
1.2 OBJETIVOS	3
1.2.1 Objetivo geral	3
1.2.2 Objetivos específicos	3
1.3 JUSTIFICATIVA E RELEVÂNCIA	3
1.4 METODOLOGIA	5
1.4.1 Procedimentos Metodológicos	6
1.5 DELIMITAÇÕES DO ESTUDO	7
1.6 COMPOSIÇÃO DO ESTUDO	8
1.7 COMENTÁRIOS GERAIS DO CAPÍTULO.....	8
2. FUNDAMENTAÇÃO TEÓRICO-EMPÍRICA	9
2.1 REVISÃO DA LITERATURA	9
2.2 CONTROLE ESTATÍSTICO DE QUALIDADE	13
2.2.1 Gráficos de controle	15
2.2.2 Gráficos de controle de Shewhart	17
2.2.3 Gráficos de controle para não-conformidades.....	18

2.3. SÉRIES TEMPORAIS	21
2.3.1 Séries estacionárias e parâmetros	21
2.3.2 Metodologia Box e Jenkins	23
2.3.3 O ciclo iterativo de modelagem	25
2.3.4 Verificação	26
2.3.5 Previsão	28
2.3.6 Critério de validação	28
2.4 MODELO DE REGRESSÃO DE POISSON	29
2.4.1 Distribuição de Poisson	29
2.4.2 Modelo de regressão de Poisson	31
2.5 MODELOS DE FUNÇÃO DE TRANSFERÊNCIA	35
2.5.1 Modelo de função de transferência de Box e Jenkins com entrada simples	38
2.5.2 Teste de causalidade	39
2.6 COMENTÁRIOS GERAIS DO CAPÍTULO.....	41
3. APRESENTAÇÃO E ANÁLISE DOS DADOS	42
3.1 COLETA DE DADOS	42
3.2 ANÁLISE DOS DADOS – ANO 2001	47
3.2.1 Análise da série gramatura da fita de polipropileno – ano 2001	47
3.2.2 Análise da série não-conforme no setor de tecelagem – ano 2001	55
3.2.3 Ajuste final – ano 2001	62
3.2.4 Comentários da análise dos dados – ano 2001	65
3.3 ANÁLISE DOS DADOS – ANO 2002	66
3.3.1 Análise da série gramatura da fita de polipropileno – ano 2002.....	66

3.3.2 Análise da série não-conforme no setor de tecelagem – ano 2002	73
3.3.3 Ajuste final – ano 2002	79
3.3.4 Comentários da análise dos dados – ano 2002	81
3.4 COMENTÁRIOS GERAIS DO CAPÍTULO	81
4. CONCLUSÕES E RECOMENDAÇÕES	83
4.1 SUGESTÕES PARA TRABALHOS FUTUROS.....	85
5. REFERÊNCIAS BIBLIOGRÁFICAS	86
ANEXO A – Modelos auto-regressivos	93
ANEXO B –Ajuste do modelo de regressão Poisson através do algoritmo de Newton-Raphson	97
ANEXO C – Função desvio	98
ANEXO D – Normas	100
ANEXO E – Relação dos dados usados no trabalho	164

LISTA DE FIGURAS

FIGURA 1 – Roteiro metodológico	6
FIGURA 2 – Gráfico de controle de Shewhart	18
FIGURA 3 – O ciclo iterativo de modelagem Box e Jenkins	26
FIGURA 4 – Gráfico representativo das distribuições para três valores de λ	30
FIGURA 5 – Fluxograma do processo	45
FIGURA 6 – Leitura diária dos dados	48
FIGURA 7 – Coeficientes da função de autocorrelação	49
FIGURA 8 – Coeficientes da função de autocorrelação parcial	50
FIGURA 9 – Resíduos da gramatura da fita de rafia	51
FIGURA 10 – Coeficientes de autocorrelação dos dados transformados.....	51
FIGURA 11 – Coeficientes de autocorrelação parcial dos dados transformados	52
FIGURA 12 – Gráfico representativo da probabilidade dos resíduos	52
FIGURA 13 – Histograma dos dados transformados	53
FIGURA 14 – Gráficos da \bar{X} e da R para os dados observados	54
FIGURA 15 – Gráficos da \bar{X} e da R para os dados transformados	55
FIGURA 16 – Contagem diária dos dados	56
FIGURA 17 – Teste de aderência.....	56
FIGURA 18 - Probabilidade half-normal dos desvios residuais	59
FIGURA 19 – Gráfico dos resíduos do desvio (G^2)	60
FIGURA 20 – Gráfico dos resíduos de Pearson	61
FIGURA 21 – Gráfico U para os dados observados	61

FIGURA 22 – Gráfico U para os dados modelados	62
FIGURA 23 – Gráfico da função de autocorrelação	63
FIGURA 24 – Gráfico da função de autocorrelação parcial.....	63
FIGURA 25 – Gráfico (U) para os dados obtidos com os resíduos da função de transferência	64
FIGURA 26 – Gráfico (U) para os dados obtidos com os resíduos da função de transferência, sem a amostra 40	65
FIGURA 27 – Leitura diária dos dados – ano 2002.....	67
FIGURA 28 – Coeficientes da função de autocorrelação.....	68
FIGURA 29 – Coeficientes da função de autocorrelação parcial.....	68
FIGURA 30 – Resíduos da gramatura da fita de rafia.....	69
FIGURA 31 – Coeficientes de autocorrelação dos dados transformados.....	70
FIGURA 32 – Coeficientes de autocorrelação parcial dos dados transformados.....	70
FIGURA 33 – Gráfico representativo da probabilidade dos resíduos	71
FIGURA 34 – Histograma dos dados transformados	71
FIGURA 35 – Gráficos da \bar{X} e da R para os dados observados	72
FIGURA 36 – Gráficos da \bar{X} e da R para os dados transformados	73
FIGURA 37 – Contagem diária dos dados	74
FIGURA 38 – Teste de aderência.....	75
FIGURA 39 - Probabilidade half normal dos desvios residuais.....	76
FIGURA 40 – Gráfico dos resíduos do desvio (G^2).....	77
FIGURA 41 – Gráfico dos resíduos de Pearson.....	77
FIGURA 42 – Gráfico U para os dados observados.....	78
FIGURA 43 – Gráfico U para os dados modelados.....	78

FIGURA 44 – Gráfico da função de autocorrelação.....	79
FIGURA 45 – Gráfico da função de autocorrelação parcial.....	80
FIGURA 46 – Gráfico (U) para os dados obtidos com os resíduos da função de transferência.....	81
FIGURA A1 – Região admissível para o ARMA (1,1) em função dos parâmetros	95
FIGURA A2 – Região admissível para o ARMA (1,1) em função das autocorrelações ρ_1 e ρ_2	96

LISTA DE TABELAS

TABELA 1 – Valores da distribuição de Poisson	30
TABELA 2 – Valores diários da gramatura da fita coletados no período de 31 de julho a 30 de setembro de 2001.....	48
TABELA 3 – Teste do qui-quadrado	53
TABELA 4 – Valores diários coletados do número de itens não conformes no setor de tecelagem no período de 01 de agosto a 31 de setembro de 2001	55
TABELA 5 – Sumário dos parâmetros do modelo	57
TABELA 6 – Critérios de avaliação do modelo	57
TABELA 7 – Estimativas do modelo	58
TABELA 8 – Critérios de avaliação do modelo	58
TABELA 9 – Sumário dos parâmetros do modelo.....	60
TABELA 10 – Critérios de avaliação do modelo	60
TABELA 11 – Critérios de avaliação do modelo.....	64
TABELA 12 - Valores diários da gramatura da fita coletados no período de 28 de fevereiro a 30 de abril de 2002	66
TABELA 13 - Teste do qui-quadrado	72
TABELA 14 - Valores diários coletados do número de itens não conformes no setor de tecelagem no período de 01 de março a 30 de abril de 2002.....	74
TABELA 15 – Sumário dos parâmetros do modelo	75
TABELA 16 – Critérios de avaliação do modelo	75
TABELA 17 – Critérios de avaliação do modelo	80
TABELA E1 – Cálculo da autocorrelação dos valores da gramatura da fita de polipropileno – ano 2001	164

TABELA E2 – Cálculo da autocorrelação dos valores da gramatura da fita de polipropileno – ano 2002	165
--	-----

LISTA DE SÍMBOLOS E ABREVIATURAS

CEQ – Controle Estatístico de Qualidade

EWMA – Médias Móveis Exponencialmente Ponderadas

CUSUM – Somas Cumulativas

i.i.d – Independentes e Identicamente Distribuídas

INAR – Modelos Autorregressivos Integrados

LSC – Limite Superior de Controle

LIC – Limite Inferior de Controle

ACF – Função de Autocorrelação

PACF – Função de Autocorrelação Parcial

MAPE – Erro Absoluto Médio Percentual

GLM – Modelos Lineares Generalizados

MARMA – Autorregressivo – Média – Móveis – Mutiplicativo

ARMA – Autorregressivo – Média – Móveis

ARIMA – Autorregressivo – Integrado – Média – Móveis

AFS – Sistema Automático de Previsão

CEP – Controle Estatístico do Processo

\bar{X} – Média Amostral

μ – Média Populacional

σ – Desvio Padrão Populacional

S – Desvio Padrão Amostral

R – Amplitude

λ – Parâmetro da Distribuição de Poisson

Z – Número de Não Conformidades

γ_k – Função de Autocorrelação de ordem k

γ_k – Autocovariância de ordem k

f_{kk} – Função de Autocorrelação Parcial

R_k – Matriz de Autocorrelação

\hat{a}_t – Resíduos Estimados

$s_k^2(\hat{a})$ – Variância Populacional dos Resíduos

p – Ordem do modelo AR

q – Ordem do modelo MA

B – Operador de Defasagem

$f(B)$ – Parâmetro do modelo AR

$q(B)$ – Parâmetro do modelo MA

$\log L$ – Função Log-verossimilhança

\mathbf{b} – Vetor de Parâmetros Desconhecidos

$(X^T W X)^{-1}$ – Matriz de Covariância Estimada

G^2 – Função Desvio

$\mathbf{c}_{n-p}^2(\mathbf{a})$ – Distribuição de Qui-quadrado

r_i – Resíduos de Pearson

\mathbf{c}^2 – Teste de Pearson

\mathbf{h}_t – Termo do Erro Normalmente Distribuído

$\nu(B)$ – Função de Transparência de Filtro

$\rho_{xy}(k)$ – Correlação Cruzada

DEDICATÓRIA

Aos meus familiares, meu muito obrigada!

A minha filha Stephanie, pela paciência de não ter a mãe presente quando precisou, pela tolerância de ser muitas vezes trocada pelo trabalho.

Ao meu esposo, parceiro de todas as horas, pela tranquilidade, carinho e compreensão que acompanharam seu apoio!

AGRADECIMENTOS

... Eu, todavia, não me esquecerei de ti. Eis que nas palmas das minhas mãos te tenho gravado.

Isaias 49.16.

- Ao meu orientador, Prof Robert Wayne Samohyl, agradeço a generosidade, compreensão e firmeza com que norteou meus estudos e pesquisas para elaboração desta tese.
- À Profa. Maria Emília Camargo, meu especial agradecimento pelas muitas horas a mim dedicadas, a co-orientação constante e segura.
- Agradeço aos membros da banca Prof. Paulo José Ogliari, Prof. Fernando Menezes Campello de Souza e, Prof. João Neiva Figueiredo, que muito contribuíram com suas sugestões
- À Indústria Têxtil Oeste Ltda., agradeço a colaboração logística.
- Aos meus colegas de estudos, a amizade e colaboração na interação dos conhecimentos.
- Aqueles que se comprometeram em ensinar-me, meus professores do curso, dedico atenção especial.
- Agradeço ao PPGE/UFSC pela oportunidade que me foi dada.
- À Universidade Regional Integrada do Alto Uruguai e das Missões – URI, especialmente o GEP, pelo suporte, apoio e incentivo na realização desta tese.

... e abrindo seus tesouros, entregaram-lhe suas ofertas: ouro incenso e mirra.

Mateus 2. 11.

I - INTRODUÇÃO

1.1 TEMA E PROBLEMA

O aumento da competitividade é o grande desafio a ser enfrentado pelas empresas brasileiras. A inevitável abertura da economia, fortalecendo a tendência a uma crescente globalização, poderá ameaçar aquelas que não admitirem esse fato. A maior competitividade está associada à melhoria da qualidade, mediante melhor controle dos processos produtivos, de forma a reduzir desperdícios e utilizar a plena capacidade dos recursos. Portanto, é de fundamental importância a adoção de formas inovadoras para o controle de qualidade, e ter esta consciência de que a busca da qualidade é um processo dinâmico, exigindo um constante aprimoramento, e avanços permanentes, assim, justifica-se a aplicação de controle estatístico em processos produtivos.

Técnicas tradicionais de controle estatístico de qualidade (CEQ), tais como Gráficos de Controle de Shewhart, Médias Móveis Exponencialmente Ponderadas (EWMA) e Somas Cumulativas (CUSUM), encontram um amplo uso no meio industrial em virtude de suas facilidades computacionais e por exigir um conhecimento mínimo de estatística. A maioria destas técnicas de controle estatístico de qualidade (CEQ), usada para detectar variações no meio de um processo, assumem que as observações não estão correlacionadas.

A metodologia tradicional de controle estatístico de qualidade é baseada em uma suposição fundamental, de que as observações são independentes e identicamente distribuídas (i.i.d), entretanto, os dados nem sempre são independentes. Quando um processo segue um modelo adaptável, ou quando o processo é uma função determinística, os dados podem ser autocorrelacionados (Johnson e Bagshaw, 1974; Harris e Ross, 1991).

Atualmente, o fenômeno da autocorrelação nas observações tem sido encontrado em um número cada vez mais expressivo nos processos. A autocorrelação nada mais é do que um mecanismo existente no processo, que faz

com que os dados não sejam independentes entre si ao longo do tempo. Em qualquer momento durante um processo, o valor de uma variável não é só um valor aleatório. Normalmente, é influenciado pelo seu próprio valor em algum momento no tempo. Isto resulta em alarmes falsos nos gráficos de controle quando não é considerado.

Sendo assim, traçar o processo de dados é extremamente valioso, no entanto, sob tais circunstâncias, não há nenhuma razão científica para usar as técnicas tradicionais de controle estatístico de qualidade, em virtude de induzir a conclusões errôneas e facilitar a uma falta de segurança de que o processo esteja sob controle estatístico com falha na identificação de variação sistemática do processo (Box e Luceno, 1997; Montgomery, 1997).

Ao se analisar variáveis contínuas, um dos procedimentos para se monitorar dados com autocorrelação, sugerido por Alwan e Roberts (1998), é usar gráficos de controle para os resíduos obtidos através de modelos de séries temporais, porém ao se tratar com variáveis discretas (número de não-conformidades num processo de produção), a distribuição de Poisson é muitas vezes empregada. Em um conjunto de dados discretos longitudinais no qual a variável resposta não pode ser transformada numa distribuição normal, estes dados são mais difíceis de se analisar. Muitos estudos têm se preocupado em encontrar a regularidade dos dados, assumindo ser um processo Poisson e, freqüentemente, são analisados através dos modelos de regressão de Poisson (Böckenholt, 1999).

Os modelos de regressão de Poisson não levam em consideração uma possível dependência entre as observações, entretanto, para dados coletados no tempo é necessário levar em consideração a dependência entre as observações, para se obter uma correta análise estatística (Zeger e Liang, 1986).

Assim, o tema aqui proposto é investigar o desempenho e a adequação do uso tradicional dos métodos de controle estatístico em processos não-estacionários, e discutir o uso da metodologia de Séries Temporais para modelar variáveis contínuas correlacionadas e o uso de modelos de Regressão de Poisson para modelar variáveis discretas correlacionadas.

1.2 OBJETIVOS

1.2.1 Objetivo geral

Apresentar uma metodologia, através dos gráficos de controle modificados, para análise de variáveis não-conformes autocorrelacionadas de um sistema produtivo.

1.2.1 Objetivos específicos

- Levantar metodologias que servirão de suporte para a análise dos dados;
- Desenvolver uma metodologia para encontrar as possíveis causas da falta de controle no processo final da produção;
- Investigar a dependência serial em dados de contagem dos itens não-conformes, utilizando os modelos de regressão de Poisson;
- Investigar a dependência serial em variáveis contínuas não-conformes, utilizando-se os modelos de séries temporais;
- Verificar o efeito da autocorrelação nos gráficos de controle;
- Apresentar, de forma sistemática, um modelo estocástico que descreva o tipo de relação entre as variáveis discretas e contínuas;
- Sintetizar os procedimentos utilizados para a implantação desta ferramenta;
- Aplicar a metodologia na Indústria Têxtil Oeste Ltda.

1.3 JUSTIFICATIVA E RELEVÂNCIA

As técnicas alternativas baseadas na metodologia de séries temporais podem ser aplicadas a certas condições previstas em controle de processos produtivos, principalmente, porque normalmente as observações não são independentes e consideram a distribuição dos dados que estão sendo analisados num processo produtivo, bem como questionar o caráter estático e/ou dinâmico com que os dados serão considerados. A dinâmica do processo é uma característica altamente sugestiva, dado que um processo produtivo fica sujeito às variações que podem facilmente alterar os patamares da produção (Montgomery, 1997).

A característica central dos trabalhos sobre CEQ, visando a melhorar o monitoramento estatístico de variáveis, tem sido a realização de estudos comparativos da capacidade preditiva destas com relação aos modelos convencionais. Quando as técnicas tradicionais de controle estatístico de qualidade são utilizadas, a possibilidade de detectar um item defeituoso só é possível após o produto ser processado, sem que se possa tomar uma medida corretiva antes da sua elaboração. Para diminuir a ocorrência de itens defeituosos no final do processo, propõe-se verificar o comportamento produtivo no processo anterior, visualizando-se, assim, o comportamento futuro dessas variáveis.

Desta maneira, será possível intervir no sistema, evitando a produção final de itens defeituosos, sempre que haja no setor anterior um afastamento das variáveis nos limites estabelecidos. A precaução de um defeito, baseada no erro de saída de uma variável, é definida como controle, pois os valores das próximas variáveis servem para alimentar o processo posterior.

Considera-se de caráter relevante este estudo, na área científica, por ser inovador na aplicação de técnicas de controle de qualidade em observações autocorrelacionadas e, o diferencial deste trabalho, que inova na literatura, consiste em modelar variáveis discretas autocorrelacionadas, utilizando-se dos modelos de Regressão de Poisson para a melhoria da qualidade em CEQ.

Considera-se de caráter relevante este estudo, no setor industrial, por ser uma forma de se fazer uma monitoração contínua, possibilitando uma melhoria no controle do sistema, dessa maneira a equipe de produção pode detectar onde um futuro problema poderá ocorrer, ou seja, localizadas as possíveis causas de instabilidade no processo, pode-se controlar a etapa seguinte, de maneira a se conseguir o menor número possível de não conformidade no resultado do processo.

Assim, justifica-se a proposta de realizar um estudo do Controle Estatístico de Qualidade, utilizando-se dos modelos de séries temporais para o tratamento de variáveis contínuas de não-conformidade autocorrelacionadas, e a apresentação e discussão da metodologia de Regressão de Poisson para o tratamento de variáveis discretas de não-conformidade autocorrelacionadas.

1.4 METODOLOGIA

A metodologia utilizada foi à pesquisa teórico-empírica (primeiro discussão pela teoria, depois diagnóstico), usando a análise descritiva, exploratória e explicativa. A pesquisa constou, então, de um embasamento bibliográfico, para levantar as noções teóricas a respeito da metodologia estatística proposta.

Os dados foram coletados através da observação simples, periódica e consulta a documentos organizacionais. A seguir, descreve-se a população estudada, o local de coleta dos dados, para um melhor entendimento do contexto onde a decisão será analisada.

Foi analisada a Indústria Têxtil Oeste Ltda., cuja implantação do CEP se deu em 1999. As visitas à empresa foram realizadas desde janeiro de 2000 e foram até maio de 2002. Os dados utilizados são diários, coletados nos três turnos em que a indústria trabalhou. Não foram utilizados formulários padrão ou outro método de pesquisa de campo, no entanto, durante as visitas técnicas procurou-se o maior número de situações práticas possíveis, seja através das observações ou de relatos das pessoas envolvidas com a implantação do CEP.

As ferramentas auxiliares utilizadas, para a análise dos dados, foram os pacotes computacionais *Statistica* e o *AutoBox*.

O pacote computacional *Statistica* foi utilizado para a análise exploratória dos dados, para a modelagem Box e Jenkins e para a modelagem de Regressão de Poisson. O *Statistica* foi desenvolvido pela Statsoft e, serve para a análise estatística em micro-computadores do tipo PC, trata de um sistema interativo, dirigido através de menus conectados hierarquicamente. Em qualquer estágio de uma sessão, se tornam disponíveis ao usuário várias facilidades que podem ser executadas através da seleção de um menu apropriado. Além de análises estatísticas, possui uma variedade de atividades com os dados. Possui um menu bem explicativo e é de fácil uso e possui, também, uma fácil importação dos dados.

O pacote computacional *AutoBox* foi utilizado para a análise do modelo via função de transferência. O Sistema Automático de Previsão (AFS) é um programa computacional, construído por David Reilly em 1975, o qual, além de modelar séries

univariadas, através dos modelos ARIMA de Box e Jenkins, também modela função de transferência, de forma automática ou não.

1.4.1 Procedimentos metodológicos

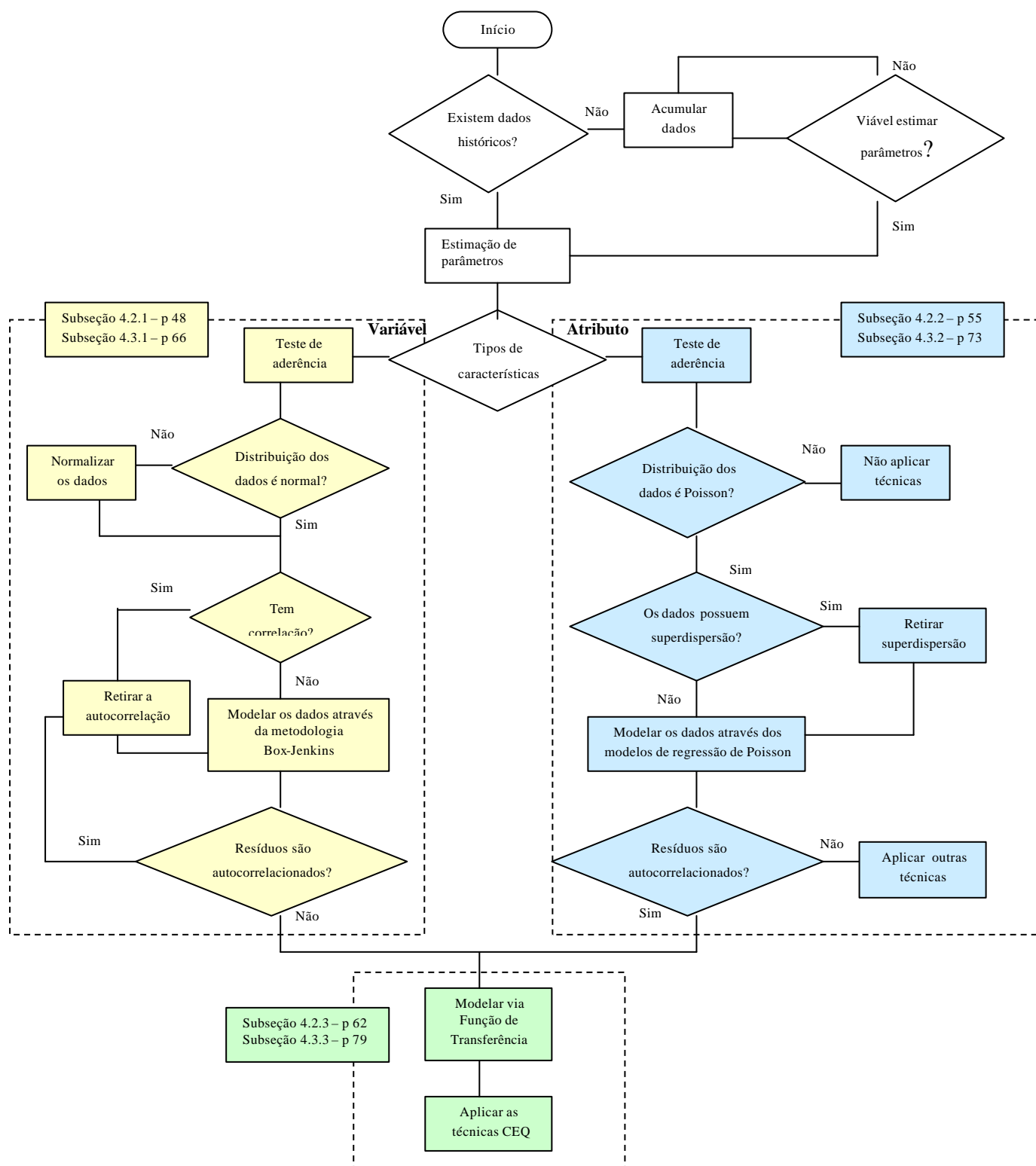


FIGURA 01 - Roteiro metodológico

O fluxograma da figura 1, página 6, apresenta o roteiro metodológico proposto a ser empregado na análise dos dados da Indústria Têxtil Oeste Ltda., sendo um procedimento referencial para outras análises em outras empresas.

Para o desenvolvimento do fluxograma esboçado na figura 1, página 6, desenvolve-se uma revisão da literatura de séries temporais, da literatura de regressão de Poisson e, da literatura de Controle Estatístico de Qualidade, procurando-se chegar ao estado da arte. Os estudos destes temas permitem o desenvolver de um novo procedimento, alcançando-se, assim, o objetivo geral proposto na página 3.

Com os dados coletados, analisou-se as variáveis contínuas através da metodologia de Box e Jenkins, e, analisou-se as variáveis discretas através dos modelos de regressão de Poisson. Com os resíduos obtidos, modelou-se as observações via função de transferência e, após, aplicou-se as técnicas de Controle Estatístico de Qualidade.

1.5 DELIMITAÇÃO DO ESTUDO

Se de um lado a carência de literatura específica, que trata de modelagem de variáveis discretas autocorrelacionadas, é o fator positivo à contribuição desta proposta, por outro lado foi um dos maiores obstáculos para a realização deste trabalho. A maior parte da literatura é direcionada às questões da área da saúde, tais como os artigos de Ferrari et ali (2002), Braga et ali (2001), Conceição et ali (2001), Wong et ali (2001), Zahner e Daskalakis (1997); não encontrando nenhum artigo voltado à indústria.

A delimitação do tipo de pesquisa proposta, também, depende da capacidade de avaliação e de constatação do pesquisador, estando associado à dificuldade de se obter informações precisas e completas, assim, pretende-se apresentar um conjunto de diretrizes para nortear uma implantação de CEP.

Neste trabalho, restringe-se a analisar os gráficos de controle para variáveis contínuas e discretas, como ferramentas utilizadas para o controle do processo. Os gráficos de controle convencionais de Shewhart, foram utilizados acrescidos de

outros modelos adequados a transformações de observações autocorrelacionadas em observações que sejam independentes e normalmente distribuídas.

Não se inclui, também, aqui uma análise dos custos dos diferentes níveis dos fatores controláveis. Porém, considera-se que, melhorando a qualidade do produto ou processo, estão se reduzindo os custos devido à má qualidade. O trabalho também não se preocupa em verificar as vantagens (ou desvantagens) das técnicas propostas como instrumentos de identificação dos efeitos significativos, ou seja, sob o enfoque de testes de hipóteses, ou da eficiência relativa. Também não se buscam soluções analíticas, pois os modelos matemáticos e estatísticos deste trabalho são demasiadamente complexos.

1.6 COMPOSIÇÃO DO ESTUDO

O presente trabalho está dividido em 5 capítulos. No capítulo 1, introduz-se o assunto a ser abordado, as razões que tornam importante a realização da pesquisa, bem como a apresentação do problema e os objetivos. Ainda, no capítulo 1, mostra-se a metodologia empregada, apresentando os métodos e técnicas que proporcionam as bases lógicas para o desenrolar da análise. No capítulo 2, é feita uma revisão dos principais estudos bibliográficos, mostrando o que se tem feito e publicado correlatamente a esse trabalho. O capítulo 3 apresenta os dados a serem utilizados e a parte dos resultados e análises das séries. No capítulo 4, conclui-se o trabalho e fazem-se recomendações para pesquisas futuras na área. O capítulo 5 dispõe do referencial bibliográfico utilizado no estudo. Em anexo, estão as demonstrações, Normas e tabelas necessárias no desenrolar do trabalho.

1.7 COMENTÁRIOS GERAIS DO CAPÍTULO

Neste capítulo estão as linhas em que a pesquisa será desenvolvida, o objetivo e as justificativas. Foi apresentada a metodologia utilizada, as ferramentas adotadas para a operacionalização da pesquisa e suas delimitações. No capítulo seguinte, será apresentada a revisão da literatura, onde se aborda como trabalhos relevantes foram conduzidos e como poderão dar suporte no desenvolver deste trabalho.

II – FUNDAMENTAÇÃO TEÓRICO-EMPÍRICA

Neste capítulo, apresenta-se a fundamentação teórico-empírica, dividida em seções, que servirá de suporte para o desenvolvimento do trabalho ora apresentado. Na seção 2.1, aborda-se a revisão da literatura centrada nos autores que proporcionaram as maiores contribuições para o estudo. Os autores são apresentados de forma cronológica, buscando transmitir uma seqüência histórica lógica dos fatos científicos mais relevantes ao assunto tratado. Na seção 2.2, mostra-se o conceito sobre gráficos de controle de qualidade, sem dúvida, uma das mais poderosas técnicas desenvolvidas, visando a auxiliar no controle eficaz da qualidade. Na seção 2.3, será apresentada a metodologia de séries temporais que serve para identificar, ajustar e diagnosticar os modelos para dados autocorrelacionados. Na seção 2.4, mostra-se o desenvolvimento de um modelo de regressão de Poisson para a análise de séries discretas no tempo. Na seção 2.5, utiliza-se à análise da função de transferência para modelar os resíduos a serem encontrados pelo modelo de séries temporais e os resíduos a serem encontrados pelo modelo de regressão de Poisson. Desta forma, pretende-se apresentar como estas técnicas serão desenvolvidas e utilizadas no estudo.

2.1 REVISÃO DA LITERATURA

A história do Controle da Qualidade é tão antiga quanto a história da própria indústria. Antes da Revolução Industrial, a qualidade era controlada pela imensa experiência dos artesões da época, o que garantia a qualidade do produto. O sistema industrial sofreu uma nova era técnica, onde o processo de produção dividia as operações complexas em tarefas simples, que podiam ser executadas por trabalhadores com habilidades específicas. Sendo assim, o operário deixou de ser o responsável por toda a fabricação do produto, ficando com a responsabilidade apenas por uma parte dele (Juran,1993).

É dentro deste contexto que surge a inspeção, onde se visava a separar os itens não conformes, a partir do estabelecimento das especificações e dos limites de tolerância. A simples inspeção não melhorava a qualidade dos produtos, apenas

fornecia informações sobre o nível de qualidade destes e separava os itens conformes, daqueles não conformes. A preocupação constante com os custos e com a produtividade deu origem à seguinte indagação: como utilizar as informações obtidas com a inspeção para melhorar a qualidade dos produtos? (Grifo, 1997)

A solução deste questionamento levou ao reconhecimento de que a variabilidade era um fator inerente aos processos industriais e podia ser compreendida através da estatística e da probabilidade, notando-se que podiam ser feitas medições durante o processo de fabricação, sem precisar esperar a conclusão do ciclo de produção.

Em 1924, Dr. Walter. A. Shewhart, da Bell Telephone Laboratories, desenvolveu um gráfico estatístico para monitorar e controlar o processo produtivo, sendo uma das ferramentas de Controle Estatístico da Qualidade. O propósito destes gráficos era diferenciar entre as causas aleatórias¹ inevitáveis e as causas assinaláveis em um processo. Segundo Shewhart (1931), se as causas aleatórias estavam presentes, não se deve mexer no processo, se as causas assinaláveis estão presentes, deve-se detectá-las e eliminá-las. Em outras palavras, estes gráficos monitoram a variação ou a falta de instabilidade no processo desse modo asseguram produtos com qualidade.

Estudos realizados por Johnson e Basgshaw (1974) e Harris e Ross (1991) mostraram que os gráficos de Shewhart e de Somas Acumuladas (CUSUM) são sensíveis à presença de dados autocorrelacionados (dados que não sejam independentes entre si ao longo do tempo), especialmente quando a autocorrelação é extrema, ou seja, não são ferramentas adequadas para o controle do processo. Será necessário primeiro tratar os dados para depois controlá-los estatisticamente.

A presença de autocorrelação nos dados leva ao crescimento no número de alarmes falsos. Alwan e Roberts (1988) demonstram que muitos alarmes falsos (sinais de causas especiais) podem ocorrer na presença de níveis moderados de autocorrelação, sendo resultantes do sistema de mensuração, da dinâmica do processo ou ambos os aspectos, e se os gráficos de controle convencionais forem utilizados sem que se conheça a presença ou não de correlação, muito esforço pode ser despendido inutilmente.

¹ Atualmente, alguns autores usam a terminologia causas comuns ao invés de causas aleatórias e causas especiais ao invés de causas assinaláveis.

Muitos métodos têm sido propostos para tratar de dados autocorrelacionados. O interesse de estatísticos na área foi estimulado pelos trabalhos de Box e Jenkins, divulgados em 1970 na obra intitulada **Time Series Analysis: Forecasting and Control**, onde foi apresentada, entre vários métodos quantitativos, a metodologia utilizada para analisar o comportamento de variáveis de séries de tempo. A metodologia Box e Jenkins utiliza o conceito de filtro que estava, na época, em franca aceitação e são definidos com precisão os processos autorregressivos e os de médias móveis.

O uso de modelos de Séries Temporais para a construção de gráficos de controle foi abordado por Montgomery e Friedman (1989). As contribuições do artigo com respeito ao uso de Séries Temporais foram de esboçar as metodologias de Séries Temporais e indicar as semelhanças de outros métodos de previsão conhecidos na engenharia, também discutiram a importante função de dados não-estacionários nas propriedades resultantes dos gráficos de controle.

Monitorar processos autocorrelacionados utilizando-se de gráficos de controle nos resíduos das observações foi discutido por Harris e Ross (1991), onde discutiram o impacto da autocorrelação nos gráficos CUSUM – *Cumulative Sum* (Somas Cumulativas) e EWMA – *Exponential Weight Moving Average* (Médias Móveis Exponencialmente Ponderadas) e apontaram que a média destes gráficos era sensível a apresentar autocorrelação. A razão para se monitorar processos residuais é que eles são independentes e identicamente distribuídas com média zero, quando o processo está controlado e permanece independente com possíveis diferenças na média quando o processo sai fora do controle.

O estudo de Alwan (1990) mostra uma análise de aplicações de efeitos mascarados das causas especiais pela autocorrelação dos dados. Embora a autocorrelação foi reconhecida mais claramente em processos contínuos e em indústrias de alta tecnologia, há evidência de que efeitos de série de tempo aparecem difundidos pela maioria das aplicações (Montgomery e Friedman, 1989; Baxley, 1990; MacGregor e Harris, 1990; Hunter, 1990).

Para acomodar os dados da autocorrelação, vários estudos foram desenvolvidos, uma proposta foi feita por Alwan e Roberts (1988) de usar um gráfico

de controle residual, onde este procedimento requer um modelo de séries temporais para o processo de dados e obter os resíduos do processo. Seguindo a mesma linha, Montgomery (1997) resume uma estratégia baseada em gráficos de controle nos resíduos de observações modeladas através de série temporais.

Segundo Zhang (1998), os gráficos tradicionais de Shewhart, os gráficos CUSUM, os gráficos EWMA podem ser aplicados aos resíduos, pois o uso de gráficos de controle residual tem a vantagem de poderem ser aplicados a dados autocorrelacionados, mesmo se os dados forem de processos não estacionários. Quando um gráfico de controle residual é aplicado a um processo não estacionário, pode-se somente concluir que o processo tem algum desvio no sistema, porque num processo não estacionário não há média constante e/ou variância constante.

Uma distribuição de probabilidade muito usada em controle de qualidade é a distribuição de Poisson, a qual é aplicada para modelar o número de defeitos ou não conformidade que ocorrem por unidade produzida (Montgomery, 1997).

Modelos de regressão de Poisson são amplamente usados em análises de dados de contagens (Dean e Lawless, 1989). Os modelos clássicos de regressão de Poisson não levam em consideração uma possível dependência entre as observações. Cox (1981) estudou a autocorrelação em dados de contagens, caracterizando duas classes de modelos dependentes no tempo: modelos condicionados a observações passadas e modelos baseados em processos latentes. Na primeira classe, podemos considerar os modelos autorregressivos e os modelos de função de transferência, onde a distribuição condicional da série é especificada em função das observações passadas. Estudos de modelos lineares generalizados com modelos condicionados são encontrados em Zeger e Qaqish (1988). Nos modelos baseados em processos latentes, a autocorrelação é introduzida através de um processo latente, onde a série depende de um erro não observável. Um modelo baseado em processos latentes, que permite incorporar estruturas de dependências entre séries de contagens, foi sugerido por Zeger (1988).

Versões de mistura entre modelos autorregressivos integrados (INAR) com modelos de regressão de Poisson para investigar a regularidade e a previsão do

comportamento de séries temporais, é apresentado por Böckenholt (1999), onde analisa a heterogeneidade e a dependência serial em dados de contagem longitudinal. Ao contrário do modelo de Zeger (1988), a versão do modelo de regressão de Poisson INAR pode ser usada quando os efeitos das covariâncias na distribuição marginal são de interesse primário (considerando a dependência entre as observações como um ruído branco), e quando a distribuição condicionada é modelada como uma função das covariâncias.

Quando há diferença nos parâmetros, isto é, o valor da variância excede à média diz-se que há uma superdispersão, valor menor da variância do que a média, indica uma subdispersão. Séries de contagem analisadas através de modelos de regressão de Poisson muitas vezes exibem superdispersão, de forma que a suposição de igualdade entre média e variância não é válida (Wang et al, 1996).

Evidências de superdispersão ou subdispersão indicam que o modelo de Poisson não está adequado. Uma solução para este fenômeno é uma aproximação de misturas finitas (Dillon e Kumar, 1994). Em contraste com modelos de efeito aleatórios que tratam de superdispersão como um fator de ruído que complica a conclusão estatística, os modelos de mistura finitas facilitam a análise da autocorrelação das covariáveis, além do mais, são derivados dos efeitos da heterogeneidade não observada nos dados (Böckenholt, 1999).

2.2 CONTROLE ESTATÍSTICO DE QUALIDADE

O controle estatístico de qualidade informa quando se deve agir ou não. Quando se age no processo, estamos nos orientando para o futuro e, com isso evitando o desperdício. Quando se age no resultado, estamos nos orientando para o passado, contamos com o desperdício e não atendemos à produção.

O controle estatístico de qualidade (CEQ) é uma técnica que consiste em analisar o processo, estabelecer padrões, comparar desempenhos, verificar e estudar desvios, buscar e implementar soluções, analisar novamente o processo após as modificações, buscando a melhor performance de máquinas e/ou pessoas (Montgomery, 1997).

Uma outra definição é dada por Triola (1999), que afirma que o CEQ é um método preventivo onde os resultados são comparados continuamente, através de dados estatísticos, identificando as tendências para variações significativas, e eliminando ou controlando estas variações com o objetivo de reduzi-las cada vez mais. Para entender melhor a técnica do controle estatístico de qualidade, é necessário ter-se em mente que a qualidade de um produto fabricado por um processo esteja, inevitavelmente sujeita à variação, e que pode ser descrita em termos de dois tipos de causa.

A **causa especial** é a causa local que só afeta uma determinada máquina, operador ou período de tempo. Está diretamente relacionada com o processo em si e pode ser resolvida, após sua detecção, na própria posição de trabalho, pela atuação do operador, mecânico ou pelo próprio grupo de supervisão.

A **causa comum** é definida como uma fonte de variação que afeta a todos os valores individuais de um processo. É resultante de diversas origens, sem que nenhuma tenha predominância sobre a outra.

Quando essas variações são significantes em relação às especificações, corre-se o risco de se ter produtos não conformes, isto é, produtos que não atendem às especificações. A eliminação de causas especiais exige uma ação local, que pode ser tomada por pessoas próximas ao processo, como, por exemplo, os operários. Já as causas comuns exigem ações sobre o sistema de trabalho, que somente podem ser tomadas pela administração, visto que o processo é em si consistente, mas, mesmo assim, incapaz de atender às especificações (Ramos, 2000).

Segundo Montgomery (1997), o maior objetivo do Controle Estatístico de Qualidade é detectar rapidamente a causa das variações no processo e usar uma ação corretiva antes das unidades de não conformidades serem manufaturadas, salienta que o mais importante é a eliminação da variabilidade no processo.

Um processo está **sob controle estatístico** (ou se trata de um **processo estável**) quando as causas especiais de variação são eliminadas do processo, e os pontos plotados no gráfico de controle permanecerem dentro dos limites de controle (Montgomery, 1997).

2.2.1 Gráficos de controle

Se um produto é fabricado para satisfazer as exigências de um cliente, geralmente deveria ser produzido por um processo que é estável. Mais precisamente, o processo deve ser capaz de operar com pouca variabilidade ao redor das dimensões designadas nas características de qualidade dos produtos. Logo, como na natureza não existem dois exemplares exatamente iguais da mesma coisa, há alguma variabilidade em toda parte, mesmo onde, aparentemente, só existe identidade, como nas linhas de produção das indústrias (Montgomery, 1997).

Os gráficos de controle servem para examinar se o processo está ou não sob controle, ou seja, indicar se somente causas comuns estão atuando sobre este processo. Sintetizam um amplo conjunto de dados, usando métodos estatísticos para observar a variabilidade dentro do processo, baseado em dados de amostragem. Informam em determinado tempo como o processo está se comportando, se ele está dentro dos limites preestabelecidos, sinalizando, assim, a necessidade de procurar a causa de variação, mas não mostrando como eliminá-la (Ryan, 1989).

Foi W. A. Shewhart (1931) que introduziu os gráficos de controle em 1924, com a intenção de eliminar variações, diferenciando-as entre as *causas comuns* e *causas especiais*.

Um gráfico de controle consiste de três linhas paralelas: uma linha média que reflete o nível de operação do processo, e duas linhas externas denominadas *limite superior de controle* (LSC) e *limite inferior de controle* (LIC), calculados em função do desvio padrão de alguma variável do processo (Shewhart, 1931).

A presença de causas comuns no processo é sugerida pela ocorrência de diferenças significativas entre o valor observado e a média do processo. Supondo que os dados possuem distribuição de probabilidade normal, e se estamos interessados em controlar o processo da média, e os limites de controle são $\bar{m} \pm 3s_{\bar{x}}$, a probabilidade de um ponto cair fora dos limites é 0,0027 (0,00135 em cada lado), ou seja, se X tem uma distribuição normal, tem-se a chance de 27 em 10.000 dos valores observados caírem fora dos limites estabelecidos (Marquardt, 1984).

Sendo assim, temos:

$$P(\bar{X} < m - 3s_{\bar{x}}) = P(\bar{X} > m + 3s_{\bar{x}}) = 0,00135 \quad (2.1)$$

ou ainda

$$P(m - 3s_{\bar{x}} < \bar{X} < m + 3s_{\bar{x}}) = 0,9973 \quad (2.2)$$

Neste intervalo, 99,7% dos itens distribuem-se em torno da média, ou seja, espera-se obter itens fora deste intervalo somente três vezes em cada mil observações. Entretanto, há casos onde pode não ser válida a aproximação das distribuições binomial ou de Poisson pela normal; nessas situações, devem-se estabelecer limites de controle baseados nas reais distribuições de probabilidade dos dados (binomial ou de Poisson) (Montgomery, 1997).

Existem vários tipos de gráficos de controle, conforme os valores característicos ou finalidade, podendo-se dividi-los em gráficos de controle por atributo e gráficos de controle por variável.

Gráficos de controle por atributos – Existem duas grandes categorias de gráficos de controle para atributos: aqueles que classificam os itens em conformes ou não-conformes, como é o caso dos gráficos da fração defeituosa ou do número de defeituosos, e aqueles que consideram o número (quantidade) de não conformidades existentes, tais como os gráficos do número de defeitos na amostra ou por unidade. Segundo Ramos (1995), as dificuldades são:

- a) em função do tamanho reduzido dos lotes, a aproximação das distribuições binomial e de Poisson pela distribuição normal poderá não mais ser válida, caso em que os limites dos gráficos de controle não podem ser determinados pelas fórmulas convencionais;
- b) as distribuições de probabilidade Binomial e de Poisson podem não representar satisfatoriamente o fenômeno estudado. Isto ocorre quando as peças são fabricadas simultaneamente (moldes com múltiplas cavidades, por exemplo), em que a incidência de defeituosos ou defeitos não é independente, estatisticamente falando.

Gráficos de controle por variável – Não podem ser utilizadas para características de qualidade que não podem ser medidas, pois o controle do processo exige a monitoração da média e da variabilidade das medidas. Os gráficos de controle para variáveis utilizam dados que podem ser medidos ou que sofrem uma variação contínua.

Algumas das metodologias indicadas para a construção de diferentes gráficos de controle são o Gráfico de Shewhart, o Gráfico MOSUM – *Moving Sum*, o Gráfico EWMA – *Exponential Weight Moving Average* (Médias Móveis Exponencialmente Ponderadas) e o Gráfico CUSUM – *Cumulative Sum* (Somas Cumulativas).

2.2.2 Gráficos de controle de Shewhart

O primeiro modelo formal de gráfico de controle foi proposto pelo Dr Walter A. Shewhart (1931), que atualmente leva o seu nome. Considere X uma estatística amostral que mede uma característica do processo usado para controlar uma linha de produção. Suponha que a média populacional de X seja μ e o desvio padrão populacional seja σ . As seguintes equações são usadas para descrever os três parâmetros que caracterizam os gráficos de controle de Shewhart (Messina, 1987)

$$LSC = \mu + k\sigma_{\bar{x}} \quad (2.3)$$

$$LC = \mu \quad (2.4)$$

$$LIC = \mu - k\sigma_{\bar{x}} \quad (2.5)$$

onde LSC é o limite superior de controle, LC é a linha central ou a média do processo, LIS é o limite inferior do processo, e k é a distância dos limites de controle até a linha central, a qual é expressa como um múltiplo do desvio padrão σ . O valor de k mais usado é 3.

A figura 2, página 18, representa um exemplo de gráfico de controle de Shewhart.

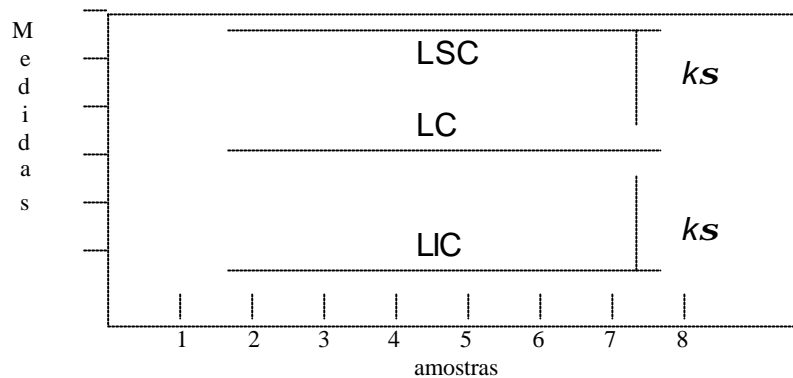


FIGURA 02 – Gráfico de controle de Shewhart

2.2.3 Gráficos de controle para não-conformidades

Um artigo não-conforme é uma unidade de produto que não satisfaz uma ou mais das especificações para aquele produto. Cada ponto determinado, no qual uma especificação não está satisfeita, resulta em um defeito ou não-conformidade. Por conseguinte, um artigo não-conforme conterá pelo menos uma não-conformidade. Porém, dependendo da natureza e severidade, é possível para uma unidade conter várias não-conformidades e não ter sido classificado como não-conforme.

É possível desenvolver gráficos de controle para o número total de não-conformidades em uma unidade ou a média do número de não-conformidades por unidade. Os gráficos de controle, normalmente, assumem que as ocorrências de não-conformidades em amostras de tamanho constantes são bem modeladas pela distribuição de Poisson. Essencialmente, isto requer que o número de oportunidades ou o potencial de localizações para não-conformidades são infinitamente grandes e que a probabilidade de ocorrência de uma não-conformidade a qualquer localização é pequena e constante. Além disso, a unidade de inspeção deve ser a mesma para cada amostra. Quer dizer, cada unidade de inspeção sempre tem que representar uma área idêntica de oportunidade para a ocorrência de não-conformidade. Além disso, nós podemos contar as não-conformidades de vários tipos diferentes em uma unidade, contanto que as condições acima estejam satisfeitas para cada classe de não-conformidades. (Messina, 1987)

Na maioria das situações práticas, estas condições não serão exatamente satisfeitas. O número de oportunidades para a ocorrência de não-conformidades pode ser finito, ou a probabilidade de ocorrência de não-conformidades pode não ser constante.

O modelo de Poisson é dado por:

$$P(Z) = \frac{e^{-\lambda} \lambda^Z}{Z!} \quad Z = 0, 1, 2, \dots \quad (2.6)$$

onde Z é o número de não-conformidades e $\lambda > 0$ é o parâmetro da distribuição de Poisson. Sabe-se que, na distribuição de Poisson, a média é igual à variância, e é igual ao parâmetro λ .

O risco de α para o limite de 3-sigmas não é alocado igualmente acima do LSC e abaixo do LIC, porque a distribuição de Poisson é assimétrica. Alguns autores recomendam o uso de limites de probabilidade quando isto ocorre, particularmente quando λ é pequeno. Então, um gráfico de controle para não-conformidades seria definido como limite de 3-sigmas. (Montgomery, 1997)

Gráficos de controle para não-conformidades:

- **Para os dados padronizados**, as equações dos limites de controle são dadas por:

$$LSC = \lambda + 3\sqrt{I} \quad (2.7)$$

$$Média = \lambda \quad (2.8)$$

$$LIC = \lambda - 3\sqrt{I} \quad (2.9)$$

assumindo que um valor padrão para λ é conhecido. Deveriam estes cálculos produzir um valor estimado negativo para o LIC, logo se fixa $LIC = 0$.

Se os dados não são padronizados, então λ pode ser estimado como a média dos números observados de não-conformidades em uma amostra preliminar de

unidades de inspeção, representado por \bar{I} . Neste caso, o quadro de controle tem parâmetros definidos como segue.

- **Para os dados não padronizados**, as equações dos limites de controle são dadas por:

$$\text{LSC} = \bar{I} + 3\sqrt{\bar{I}} \quad (2.10)$$

$$\text{Média} = \bar{I} \quad (2.11)$$

$$\text{LIC} = \bar{I} - 3\sqrt{\bar{I}} \quad (2.12)$$

Quando os dados não são padronizados, os limites de controle nas equações (2.10, 2.11 e 2.12) podem ser considerados como limites de controle do processo, e as amostras preliminares examinadas carecem de controle. O gráfico de controle para não-conformidades também é chamado às vezes de carta λ .

Modelos de probabilidade alternativos para contagem de dados

De acordo com Montgomery (1997), a maioria das aplicações dos gráficos de controle assume que a distribuição de Poisson é o modelo correto implícito no processo. Porém, não é a única distribuição que poderia ser utilizada como um modelo de contagem ou não-conformidades por dados de unidade. Vários tipos de fenômenos podem produzir itens defeituosos que não são bem modelados pela distribuição de Poisson. Por exemplo, suponha que não-conformidades tendem acontecer em agrupamentos; isto é, se há uma não-conformidade em alguma parte de um produto, então é provável que haverá outras. Note que há pelo menos dois processos aleatórios para trabalhar aqui: um que gera o número e a localização de agrupamentos e o segundo que gera o número de não-conformidades dentro de cada agrupamento. Se o número de agrupamentos tem uma distribuição de Poisson e o número de não-conformidades dentro de cada agrupamento tem uma distribuição comum, então o número total de não-conformidades tem uma combinação da distribuição de Poisson. Muitos tipos de combinações ou distribuições generalizadas poderiam ser usados como um modelo para dados do tipo contagem. Se os dados de agrupamento possuem uma distribuição gama e o número de não-conformidades dentro de cada agrupamento é Poisson, o resultado é

uma distribuição binomial negativa. Johnson e Kotz (1969) deram um bom resumo destes e de outras distribuições discretas que poderiam ser úteis modelando dados do tipo contagem.

Misturas de vários tipos de não-conformidades podem conduzir a situações em que o número total de não-conformidades não é modelado adequadamente pela distribuição de Poisson. Situações similares acontecem quando os dados da contagem têm também muitos ou poucos zeros. Uma boa discussão geral do problema é o artigo de Jackson (1972) e no livro de Hinde e Demétrio (1998). O uso da distribuição binomial negativa para modelar dados de contagem em unidades de inspeção de tamanho variado tem sido estudada por Scheaffer e Leavenworth (1976). Gardiner e Montgomery (1987) descrevem o uso de várias distribuições discretas para modelar a ocorrência de defeitos em circuitos integrados.

2.3 SÉRIES TEMPORAIS

Na presente seção, aborda-se os métodos para identificar, ajustar e diagnosticar os modelos para séries autocorrelacionadas. A correlação pode ser encontrada se a série for coletada seqüencialmente no tempo (Lago Neto, 1999).

Num primeiro momento, aborda-se sobre as séries estacionárias e seus parâmetros e a presença da autocorrelação na utilização de gráficos de controle, depois aborda-se a metodologia de Box e Jenkins, onde apresenta-se os testes para validar a consistência do modelo especificado.

2.3.1 Séries estacionárias e parâmetros

Um processo estocástico é dito estritamente estacionário se suas propriedades não são afetadas por uma mudança na origem de tempo, isto é, a distribuição conjunta $P(Z_t, Z_{t+k})$ é a mesma qualquer que seja t , onde Z é uma família de variáveis aleatórias e, t usualmente referenciado como tempo.

Sob o ponto de vista intuitivo, uma série é estacionária se :

- a) não há mudança sistemática da média (tendência);
- b) não há mudança sistemática na variância;
- c) não há variações estritamente periódicas.

A maior parte da teoria de séries temporais lida com séries estacionárias. Por isso, a série deve ser trabalhada previamente através de transformações. O gráfico da série deve mostrar as propriedades mais importantes (tendência, sazonalidade). Se a série é aproximadamente estacionária, o processo gerador estacionário pode, então, ser adequadamente descrito pelos momentos de suas distribuições de probabilidade (Murteira; Müller e Turmkman, 1993).

Os momentos incluem a média, a variância, a função de autocovariância, e sua transformada de Fourier, o espectro potência.

O estudo dos processos estacionários pode ser feito no domínio da frequência ou no domínio do tempo. O estudo no domínio da frequência dá papel de relevo aos conceitos de periodograma e de densidade espectral; o domínio no tempo atribui papel predominante às funções autocovariância e autocorrelação (Beltrão, 1991).

A autocorrelação é uma medida de dependência entre observações da mesma série separadas por um determinado intervalo chamado retardo. A função de autocorrelação (ACF) nada mais é do que a representação gráfica do coeficiente de autocorrelação em função dos diversos retardos que podem ser atribuídos aos dados (Box, Jenkins, Reinsel, 1994).

A função de autocorrelação (ACF) permite que se entenda melhor o comportamento da dependência estatística entre os dados e, posteriormente, será útil quando da determinação de qual série temporal utilizar para o ajuste do processo.

Define-se autocorrelação de ordem k como

$$r_k = \frac{g_k}{g_o} \quad (2.13)$$

estimado por:

$$r_k = \frac{C_k}{C_o} \quad (2.14)$$

onde $\mathbf{g}_k = E[(Z_t - \mathbf{m})(Z_{t+k} - \mathbf{m})]$ é a autocovariância de ordem k

$\mathbf{g}_0 = \mathbf{s}_Z^2$ é a variância para cada Z_t

$$C_k = \bar{\mathbf{g}}_k = \frac{1}{n} \sum_{t=1}^{n-k} (Z_t - \bar{Z})(Z_{t+k} - \bar{Z}) \quad (2.15)$$

n = número de observações

$$\bar{Z} = \sum_{t=1}^n \frac{Z_t}{n} \text{ é a média amostral} \quad (2.16)$$

Outra ferramenta necessária à identificação do processo é a função autocorrelação parcial, medida de autocorrelação de ordem k condicionada ao conhecimento das autocorrelações em todos os retardos de ordem inferior a k . É definida por $\{f_{kk}\}$ onde:

$$f_{kk} = \frac{|R_k^*|}{|R_k|} \quad (2.17)$$

sendo: R_k a matriz de autocorrelação já definida; $|R_k|$ seu determinante; R_k^* a matriz que difere de R_k somente pela última coluna, substituída por (r_1, r_2, \dots, r_k) (Beltrão, 1991).

2.3.2 Metodologia Box e Jenkins

A generalização dos modelos autorregressivos (Box e Jenkins) encontra-se no Anexo A.

Identificação dos modelos - Segundo Box e Jenkins (1976), dada uma série estacionária Z_1, Z_2, \dots, Z_n , pode-se estimar a média \mathbf{m} e a autocovariância de ordem k $\{\mathbf{g}_k\}$ através de:

$$\hat{\mathbf{m}} = \bar{Z} = \frac{\sum_{t=1}^n Z_t}{n} \quad (2.18)$$

e

$$\hat{g} = C_k = \frac{\sum_{t=1}^n (Z_t - \bar{Z})(Z_{t+k} - \bar{Z})}{n} \quad (2.19)$$

Assim, a função de autocorrelação r_k pode ser estimada por:

$$\hat{r}_k = r_k = \frac{C_k}{C_0} \quad (2.20)$$

Para um processo normal estacionário, Bartlett (1946), estabelece que,

$$\text{supondo-se } r_k = 0, \quad \text{Cov}(r_k, r_{k+s}) \cong \frac{\sum_{t=-k+s}^k r_t r_{t-s}}{n} \quad \text{para } s=0 \text{ tem-se } \text{Var}(r_k) \cong \frac{\sum_{t=-s}^k s_t^2}{n} \text{ e,}$$

para n suficientemente grande r_k tem distribuição normal com média nula e variância estimada por:

$$\text{Var}(r_k) \cong \frac{(1 + 2 \sum_{t=1}^k r_t^2)}{n} \quad (2.21)$$

A raiz quadrada da equação 2.21 é chamada de erro padrão de grande retardo.

Procedimentos – Inicialmente, constrói-se o gráfico da série para inspeção visual da estacionariedade. Calculam-se média, variância, autocorrelações e autocorrelações parciais $\{f_{kk}\}$.

A inspeção de $\{r_k\}$ e $\{\hat{f}_{kk}\}$ deve indicar o modelo em estudo. Observa-se a ocorrência de cortes em $\{r_k\}$ e $\{\hat{f}_{kk}\}$, comparando-se as autocorrelações estimadas com seu erro padrão de grande retardo e, então, vendo se os resultados seguem um padrão teórico dentre a classe geral de modelos autorregressivos (ARMA(p,q)) Os valores da ordem do processo (p e q) são obtidos a partir de (Box, Jenkins, Reinsel, 1994):

$$\hat{f}_{kk} \sim N(0; 1/N), \quad \forall k > p \rightarrow \text{modelo AR}(p) \quad (2.22)$$

$$r_k \sim N(0; \frac{1}{N} [1 + 2 \sum_{t=1}^q r_t^2]), \quad \forall k > q \rightarrow \text{modelo MA}(q) \quad (2.23)$$

$\{\hat{f}_{kk}\}$ e $\{r_k\}$ devem ser usados somente como indicadores genéricos, devido à alta correlação existente entre as r_k . O estágio da identificação requer computação automática e inspeção experiente das autocorrelações estimadas, ver maiores detalhes em Box e Jenkins (1976) ou Souza e Camargo (1996).

É conveniente, construir o histograma dos valores observados para dar suporte à hipótese de normalidade.

Se a média (\bar{Z}) da série é diferente de zero, deve-se trabalhar com a série desviada da média $\tilde{Z}_t = Z_t - \bar{Z}$. Assim, é preciso testar se $E(Z_t)=0$ comparando-se a média (\bar{Z}) com seu desvio padrão ($s_{\tilde{z}}$), que depende do processo. Deve-se também calcular, para cada modelo identificado, as estimativas iniciais dos parâmetros e da variância do ruído branco. O modelo que não se enquadrar nas restrições de admissibilidade - estacionariedade e inversibilidade (maiores detalhes no Anexo 1) - deverá, em princípio, ser rejeitado. Deve-se, também, testar a significância de cada parâmetro, comparando sua estimativa com o desvio associado.

2.3.3 O ciclo iterativo de modelagem

Feita a identificação do modelo genérico ARMA(p,q), os modelos AR(p) e MA(q) podem ser considerados como casos particulares do ARMA(p,q), passa-se para a obtenção das estimativas de máxima verossimilhança (eficientes) para os (p+q+2) parâmetros ($f_1, \dots, f_p, q_1, \dots, q_q, m, s_a^2$).

Em seguida, passa-se ao diagnóstico do modelo, analisando-se a série de resíduos proveniente do ajustamento. Sendo aceito o modelo como ajustado, passa-se à fase de previsão, caso contrário, a análise dos resíduos deve indicar o novo modelo corrente (Box e Jenkins, 1976).

O esquema da figura 3, página 26, mostra o ciclo iterativo da modelagem Box e Jenkins.

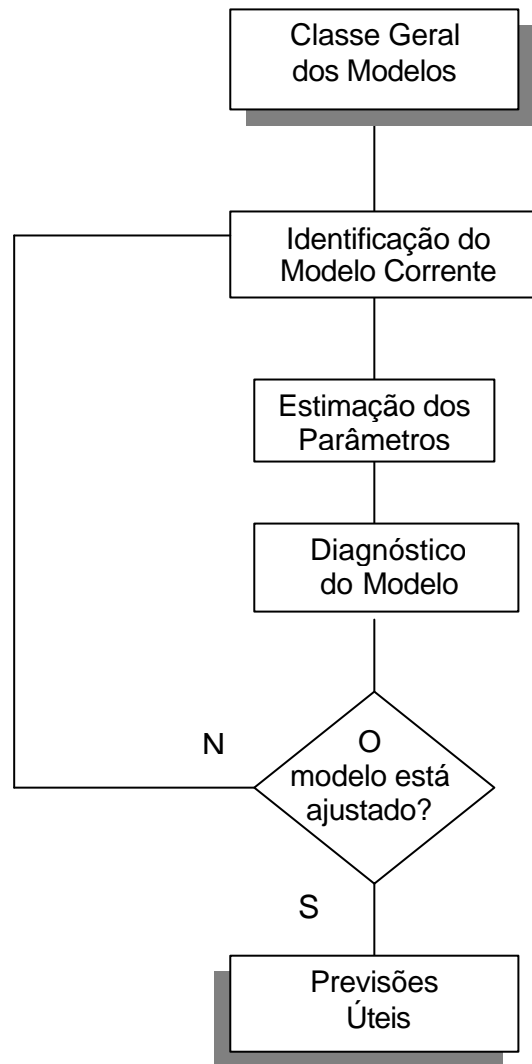


FIGURA 03 – O ciclo iterativo de modelagem Box e Jenkins

2.3.4 Verificação

Obtidas as estimativas dos parâmetros resta decidir se o modelo é adequado. Esta decisão é tomada mediante a verificação dos resíduos estimados $\hat{a}_t = \tilde{Z}_t - \hat{Z}_t$, onde \tilde{Z}_t é a série desviada e \hat{Z}_t é a série prevista, pois, se o modelo é correto tem-se:

$$r_k(\hat{a}) \text{ i.i.d } \sim N(0, \frac{1}{N}) \quad (2.24)$$

A verificação pode ser feita através do que se segue (Box e Jenkins, 1976):

1) Autocorrelação dos resíduos: inspeção do gráfico $r_k(\hat{a})$. Se o modelo é adequado, as autocorrelações $r_k(\hat{a})$ devem estar praticamente todas dentro dos limites de ± 2 desvios padrões.

2) Teste de Durbin-Watson: tem por objetivo testar a hipótese nula de que não há autocorrelação de primeiro grau nos resíduos $\{a_t\}$.

3) Teste de Portmanteau: $Q = N \sum_{k=1}^{\bar{k}} s_k^2(\hat{a})$, onde N é o número de elementos e $s_k^2(\hat{a})$ a variância populacional dos resíduos. Através da estatística Q testa-se a hipótese nula de que as autocorrelações, $r_k(\hat{a})$, $k = 1, 2, \dots, \bar{k}$ são iguais a zero. Se o modelo não é apropriado, os valores médios de Q tenderão a crescer. Q tem uma distribuição de Qui-quadrado (c_{k-p-q}^2) onde:

\bar{k} = número de autocorrelações estimadas;

p = número de parâmetros AR;

q = número de parâmetros MA.

Se a verificação do diagnóstico acusa inadequação do modelo, é necessário encontrar o novo modelo para estudo.

Suponha-se que em uma série foi ajustado um modelo AR(1), isto é,

$$a_t = Z_t - fZ_{t-1} = f(B)Z_t \quad (2.25)$$

e que a inspeção das $r_k(\hat{a})$ acusou sua inadequação. Então:

a) aplica-se à série de resíduos \hat{a}_t a metodologia Box e Jenkins;

b) identificado o modelo em a_t , por exemplo, um MA(1), temos:

$$a_t = q(B)e_t = e_t + q e_{t-1} \quad (2.26)$$

substituindo a equação 2.26 no modelo AR(1) $a_t = Z_t - fZ_{t-1} = f(B)Z_t$, equação 2.25, o novo modelo tentativo fica:

$$e_t + q e_{t-1} = Z_t - f Z_{t-1}, \text{ onde } (1 + q B)e_t = (1 - f B)Z_t, \text{ o qual resulta em:}$$

$$\mathbf{q}(B)e_t = \mathbf{f}(B)Z_t \quad \text{que será um modelo ARMA (1,1).} \quad (2.27)$$

Se vários modelos são ajustados, o melhor será o que apresentar menor valor de Q.

2.3.5 Previsão

Obtido o modelo ajustado e, conseqüentemente as estimativas dos parâmetros, parte-se para a utilização do modelo em previsão. Para uma série gerada por um ARMA ($\mathbf{f}(B)Z_t = \mathbf{q}(B)a_t$), define-se $\hat{Z}_t(\ell)$ como uma previsão de $Z_{t+\ell}$ feita no instante t para o avanço ℓ (Box e Jenkins, 1976).

A partir da origem t, far-se-á previsão $\hat{Z}_t(\ell)$ de $Z_{t+\ell}$ que pode ser uma função linear decorrente de valores prévios Z_{t-1}, Z_{t-2}, \dots , como também de resíduos $a_t, a_{t-1}, a_{t-2}, \dots$

Tomando-se a função de previsão:

$$\hat{Z}_t(\ell) = \mathbf{y}_\ell a_t + \mathbf{y}_{\ell+1} a_{t-1} + \mathbf{y}_{\ell+2} a_{t-2} + \dots \quad (2.28)$$

onde $\mathbf{y}_\ell, \mathbf{y}_{\ell+1}, \dots$ são pesos a serem determinados.

O erro no instante t, para o avanço ℓ , é dado por:

$$e_t(\ell) = Z_{t+\ell} - \hat{Z}_t(\ell) = a_{t+\ell} \quad (2.29)$$

com $\text{Var}(e_t(\ell)) = (1 + \mathbf{y}_1^2 + \mathbf{y}_2^2 + \dots + \mathbf{y}_{\ell-1}^2) \mathbf{s}_a^2$.

O intervalo de confiança para a previsão, com um nível de significância α é:

$$Z_{t+\ell} \in \left[\hat{Z}_t(\ell) \pm t_{\frac{\alpha}{2}} (\mathbf{s}_a^2 \sum_{j=0}^{\ell-1} \mathbf{y}_j^2)^{\frac{1}{2}} \right]. \quad (2.30)$$

2.3.6 Critério de validação

A título de validar o modelo ajustado, com vistas à realização das previsões, será utilizado o MAPE (Erro Absoluto Médio Percentual). O MAPE será calculado a

partir das previsões um passo à frente gerado por cada modelo estimado (Russo, 1989).

O seguinte critério de ajuste também deve ser considerado.

$$MAPE \quad (\%) = \frac{\sum \left| \frac{Z - \hat{Z}}{Z} \right|}{n}. \quad (2.31)$$

2.4 MODELO DE REGRESSÃO DE POISSON

Nesta seção, num primeiro momento, descreve-se a distribuição de Poisson e, depois, se mostra o desenvolvimento de um modelo de regressão de Poisson para a análise de séries discretas no tempo.

2.4.1 Distribuição de Poisson

O nome da distribuição de Poisson está associado ao matemático francês S. Denis Poisson (1781-1840), que a estudou em um livro onde trata da aplicação da teoria das probabilidades na justiça comum. Probabilidades Poisson são usadas para modelar o número de ocorrências de um evento. Um dos mais antigos usos da distribuição de Poisson foi em 1898 ao ser investigado o número de soldados Prussianos que eram mortos pelos coices dos cavalos (Ender, 2002).

A distribuição de Poisson é uma distribuição discreta de probabilidade, aplicável à ocorrência de um evento em um intervalo especificado. A função de distribuição de probabilidade Poisson é dada por:

$$\Pr(Z / I) = \frac{e^{-I} \cdot I^Z}{Z!} \quad \text{para } Z=0,1,2,\dots \quad \text{e } e=2,71828\dots \quad (2.32)$$

A variável aleatória Z é o número de ocorrências do evento em um intervalo. O intervalo pode ser o tempo, a distância, a área, o volume ou outra unidade análoga.

A média e a variância de uma distribuição de Poisson são iguais a λ (lambda), que é conhecido como o parâmetro da distribuição de Poisson. A tabela 1

representa as probabilidades da distribuição de Poisson para $\lambda = 1, 3$ e 5 . Após a tabela 1, segue o gráfico representativo.

TABELA 1 - Valores da distribuição de Poisson

Z	$\lambda = 1$	$\lambda = 3$	$\lambda = 5$
0	0,3678795	0,04978707	0,00673795
1	0,3678795	0,14936121	0,03368973
2	0,1839397	0,22404180	0,08422434
3	0,0613132	0,22404180	0,14037390
4	0,0153283	0,16803135	0,17546737
5	0,0030657	0,10081881	0,17546737
6	0,0005109	0,05040941	0,14622281
7	0,0000730	0,02160403	0,10444486
8	0,0000091	0,00810151	0,06527804
9	0,0000010	0,00270050	0,03625580
10	0,0000001	0,00081015	0,01813279

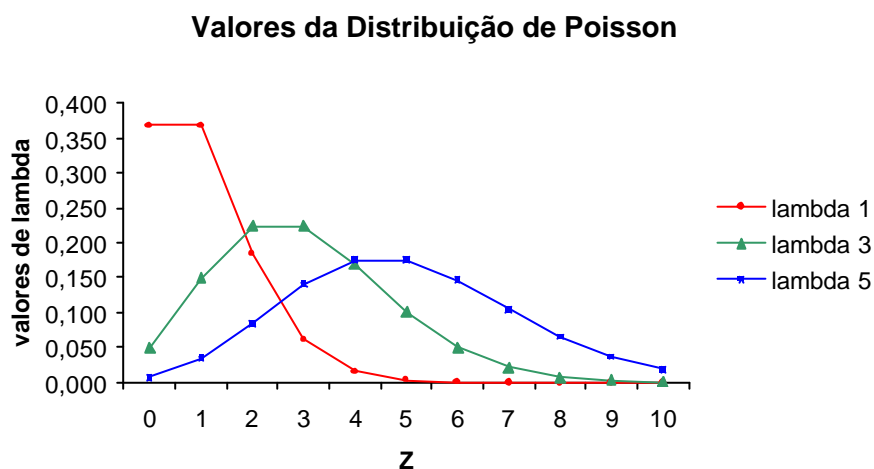


FIGURA 04 - Gráfico representativo das distribuições para três valores de λ .

Pela figura 4, observa-se ao crescimento do valor de lambda há o achatamento da curva e o deslocamento da mesma para a direita. Para valores grandes de lambda, a distribuição é aproximadamente normal.

A distribuição, na qual a média é igual à variância, é dita que possui igual dispersão. Quando a variância é maior do que a média, diz-se que há uma superdispersão. Na prática, é muito raro haver distribuição com dispersão igual.

2.4.2 Modelo de regressão de Poisson

O interesse pelo estudo de análise de dados de contagens tem tido um crescimento nos últimos 20 anos (Ender, 2002).

Nas últimas duas décadas, têm sido feitos experimentos em diversos campos das engenharias, tal como na melhora de produtos e processos. Muitos experimentos industriais envolvem variáveis que não apresentam uma distribuição normal; os Modelos Lineares Generalizados (GLM) são uma útil alternativa para os métodos tradicionais de análise dos dados que necessitam de transformações. Os GLM, introduzidos por Nelder e Wedderburn (1972), desempenham um papel importante na Estatística, uma vez que generalizam o modelo tradicional de regressão normal linear, abrindo o leque de opções para a distribuição da variável resposta e dando maior flexibilidade para a ligação entre a média e a parte sistemática do modelo (Paula, 1997).

Depois da regressão logística, o modelo de regressão de Poisson é o mais usado dos Modelos Lineares Generalizados. Os modelos de regressão de Poisson são aplicados quando a resposta é uma contagem, tal como o número de eventos ocorridos num espaço de tempo.

Segundo Cordeiro (1992), o modelo de regressão de Poisson é um tipo específico de modelos lineares generalizados (GLM), cujos parâmetros podem ser estimados usando-se o método da máxima verossimilhança. A função de verossimilhança é dada por:

$$L = \prod_{i=1}^n \Pr(Z_i / I_i) = \prod_{i=1}^n \frac{e^{-I_i} \cdot I_i^{Z_i}}{Z_i!} \quad (2.33)$$

e a função log-verossimilhança é igual a $\log L = \sum (Z_i \cdot \log(I_i) - I_i) - \sum \log(Z_i!)$

A função constante de Z , dada por $\sum \log(Z_i!)$, pode ser omitida, pois não envolve I (Christensen, 1997; Ferrari, 2002).

O modelo envolve três componentes:

- 1) O componente aleatório: representado por um conjunto de variáveis aleatórias, Z_1, Z_2, \dots, Z_n com $E(Z_i) = I_i$;
- 2) A componente sistemática: $\mathbf{b} x_i^T$; onde \mathbf{b} é o vetor de parâmetros desconhecidos e, x representa os valores das variáveis explicativas;
- 3) Função de ligação: $\log(I_i)$, que relaciona o preditor linear ao valor esperado I do vetor Z .

Assim, a componente sistemática admite a existência de uma função de ligação $\eta = \log(I_i)$ entre as médias das observações e a estrutura linear do modelo dada por $\log(I_i) = \mathbf{b} x_i^T$. Logo, o uso da função de ligação log assegura que os valores ajustados de I_i permanecem no intervalo $[0, \infty)$. A função de ligação log relaciona o preditor linear $\mathbf{b} x_i^T$ ao valor esperado I_i do vetor Z_i . O modelo de Poisson com ligação log é algumas vezes chamado de modelo log-linear (Schafer, 1997; McCullagh e Nelder, 2000).

Transformando a função de ligação log obtêm-se a seguinte expressão para a variável resposta:

$$I = e^{b_0 + b_1 x_1 + \dots + b_k x_k} \quad (2.34)$$

onde $\mathbf{b} x_i^T$ é o preditor linear, $\mathbf{b} = (b_1, \dots, b_k)^T$ é o vetor de parâmetros desconhecidos a serem estimados e $x_i = (x_{i1}, \dots, x_{ik})^T$ representa os valores de k variáveis explicativas (Paula, 1997).

As distribuições normal, binomial, Poisson, gamma, normal inversa, e binomial negativa são membros da família exponencial (Lewis, Montgomery, Myers, 2001).

Estatísticas de interesse - Depois do ajuste do GLM, pelo método da máxima verossimilhança (que pode ser feita através do processo iterativo do tipo Newton-Raphson, exposto no Anexo B), deve-se examinar as seguintes estatísticas:

- $(X^T W X)^{-1}$, a matriz de covariância estimada para $\hat{\mathbf{b}}$, onde X é a matriz do modelo e W é a função peso que se modifica no algoritmo iterativo que converge rapidamente, na maioria dos casos;

- a função de log-verossimilhança $l=l(Z, \mathbf{I}) = \sum_{i=1}^N \{Z_i \log \mathbf{I}_i - \mathbf{I}_i\}$.

Um gráfico dos resíduos padronizados versus valores ajustados, sem nenhuma tendência, é um indicativo de que a relação funcional variância/média proposta para os dados é satisfatória. Gráficos dos resíduos versus covariáveis que não estão no modelo são bastante úteis. Se nenhuma covariável adicional é necessária, então não se deve encontrar qualquer tendência nesses gráficos. Dados com erros grosseiros podem ser detectados tendo grandes resíduos, ou o modelo ajustado deve requerer mais covariáveis, por exemplo, interações de ordem superior. A inspeção gráfica é um meio poderoso de inferência nos GLMs (Piegorisch, 1998).

Conforme mostra Schafer (1997), o desvio G^2 é a estatística do teste da razão de verossimilhança para comparar o ajuste do modelo: $\log \mathbf{I}_i = \mathbf{b} x_i^T$, com o modelo saturado¹ (desvio). O modelo saturado ajusta uma média separada para cada Z_i , não fazendo menção de como são relacionadas as covariáveis. É fácil mostrar que a verossimilhança, $Z_i \log \mathbf{I}_i - \mathbf{I}_i$, chega a um máximo quando $\mathbf{I}_i = Z_i$, assim o desvio é dado por (maiores detalhes no Anexo C):

$$G^2 = 2l(Z, Z) - 2l(\mathbf{I}, Z) \quad (2.35)$$

desenvolvendo-se a equação 2.35 tem-se:

$$G^2 = 2 \sum_{i=1}^N \left\{ Z_i \log \frac{Z_i}{\mathbf{I}_i} - (Z_i - \mathbf{I}_i) \right\}. \quad (2.36)$$

Note que esta expressão (2.36) será indefinida se tivermos $Z_i=0$. É fácil ver que se $Z_i=0$, a contribuição desta observação com relação ao desvio deve ser de $2\mathbf{I}_i$, então a melhor maneira para expressar a estatística do desvio é

$$G^2 = 2 \sum_{i:Z_i=0} \left\{ Z_i \log \frac{Z_i}{\mathbf{I}_i} - (Z_i - \mathbf{I}_i) \right\} + 2 \sum_{i:Z_i=0} \mathbf{I}_i \quad (2.37)$$

¹Modelo Saturado – O GLM tem um parâmetro para cada observação, ajusta-se exatamente aos dados, isto é, as estimativas das médias são iguais às próprias observações e toda a variação é devida à componente sistemática (Paula, 1997).

Pode-se definir o desvio residual como a raiz quadrada da contribuição da i -ésima observação do desvio, multiplicado por $(Z_i - I_i)$.

Um modelo mal ajustado aos dados tem um grande desvio e um modelo bem adequado aos dados um pequeno desvio. Porém, um grande número de parâmetros significa um grau de complexidade na interpretação do modelo.

Os graus de liberdade associados ao desvio são definidos por $\nu = n - p$, onde n é o número de parâmetros do modelo saturado e p é o número de parâmetros do modelo sob investigação. Para o teste de um modelo, compara-se o valor de G^2 e seus graus de liberdade ν , com uma distribuição teórica de probabilidade. Geralmente, adota-se a distribuição de qui-quadrado. Na prática, testa-se um modelo (sem muito rigor) comparando o desvio com o valor crítico $c_{n-p}^2(\alpha)$ da distribuição de qui-quadrado à um nível de significância igual a α . Se este for maior que $c_{n-p}^2(\alpha)$, o modelo será rejeitado e, caso seja menor ou igual, aceito (Paula, 1997).

Outra medida para a verificação do melhor ajuste é o teste de Pearson, que compara a distribuição observada com a determinada pelo modelo. Os ajustes marginais do modelo de regressão de Poisson podem ser determinados calculando-se os resíduos de Pearson: $r_i = (Z_i - I_i) / (\sqrt{I_i})$. Assim, o teste de Pearson para o melhor ajuste é dado por $c^2 = \sum_{i=1}^N r_i^2$ (utiliza-se a tabela do c^2), (Wang et al, 1996).

Deve-se esperar que um modelo bem ajustado aos dados tenha um desvio próximo dos seus graus de liberdade. Se o desvio G^2 ou o c^2 excedem o valor do seus graus de liberdade, diz-se que o modelo é inadequado, podendo tratar-se de um problema de superdispersão (Wang et al, 1996).

Além destes testes de ajustes, uma variedade de testes específicos estão disponíveis para se determinar se os dados possuem superdispersão (variância maior do que a média) ou subdispersão (variância menor do que a média), tal como o teste de pontuação proposto por Dean (1992) e o teste direto desenvolvido por Mills e Seneta (1991).

Para se avaliar a existência de superdispersão, deve-se empregar um critério uniforme. Se adota que o desvio G^2 ou o teste de Pearson c^2 excedem o valor crítico $c_{n-p}^2(\alpha)$ até 10%, caso seja maior o modelo será rejeitado. O efeito desta correção é mínimo nos estimadores pontuais. Este método está baseado na generalização de Breslow (1984) para modelos de Poisson.

O objetivo da análise é descrever $I_i = E(Z_i)$ como uma função das covariáveis. Em observações independentes, o modelo log linear pode ser usado. No modelo log linear, a média é $I_i = \exp(\mathbf{b} x_i^T)$, e, se Z_i tem distribuição de Poisson, o método de verossimilhança pode ser usado para estimar \mathbf{b} . No modelo de Poisson, a $\text{var}(Z_i) = I_i$. É muito comum na prática encontrar-se $\text{var}(Z_i) > I_i$. Nestes casos, métodos de verossimilhança, os quais levam em conta uma variedade de relações entre variâncias e médias; são apropriados (McCullagh e Nelder, 2000; Zeger, 1988).

Em outras palavras, uma maneira de ver a superdispersão é considerar que para a $\text{var}(Z_i) = I_i s^2$, alguns valores de $s^2 > 0$. Para $s^2 = 1$ o modelo pode ser ajustado por dispersão de Poisson padrão, para outros valores de s^2 , o modelo é ajustado pelo método da máxima verossimilhança, ou seja, representa ajustar por verossimilhança um modelo equivalente a uma binomial negativa (gamma-poisson) (Zeger, 1988).

2.5 MODELOS DE FUNÇÃO DE TRANSFERÊNCIA

Os modelos Box e Jenkins univariados permitem fazer a análise - descrição e previsão - de uma série temporal através do conhecimento exclusivo dos seus valores. Contrariamente, aos demais modelos que são modelos não causais: em cada momento a série é apenas explicada por valores da própria série anteriores a esse momento. A existência de fenômenos aleatórios relacionados entre si conduz ao tratamento simultâneo das correspondentes séries temporais de forma que a informação contida em cada uma possa se beneficiar desse relacionamento e, portanto, melhorar e enriquecer o estudo desses fenômenos (Box e Jenkins, 1974).

A extensão da análise univariada, a uma abordagem de diversas séries cronológicas correlacionadas representa o desenvolvimento lógico do estudo para a presente seção.

Os modelos de função de transferência combinam características dos modelos univariados de Box e Jenkins com a análise de regressão linear. Por uma parte, representam uma fecunda "mistura" de metodologias não causais e causais que permite descrever uma série, que constitui o "output" (saída), como função de uma ou mais séries, que formam o "input" (entrada). Por outro lado, representam um modelo de resposta dinâmica, uma vez que a variação de uma das séries de entrada em determinado instante produz na série de saída uma resposta retardada. Tais modelos são importantes, pois permitem definir as variáveis de entrada como indicadores avançados ("leading indicators") do "output" (Murteira; Müller e Turmkman, 1993).

Segundo Pindyck e Rubinfeld (1991), pode-se combinar análise de séries temporais e análise de regressão para produzir uma previsão melhor que o uso de qualquer um dessas técnicas sozinhas. Ao se prever a variável Z_t , usando-se um modelo de regressão, possivelmente tal modelo incluiria todas as variáveis independentes que poderiam prover uma explicação para movimentos em Z_t , mas que não sejam colineares. Por exemplo, suponha que o melhor modelo de regressão que contém duas variáveis independentes, x_1 e x_2 , seja:

$$Z_t = v_0 + v_1 x_{1Z} + v_2 x_{2Z} + e_t. \quad (2.38)$$

Esta equação tem um termo de erro aditivo implícito que responde por uma variância inexplicada em Z_t ; quer dizer, responde por aquela parte da variância de Z_t que não é explicado por x_1 e x_2 . A equação pode ser estimada, e terá um R^2 , o qual (a menos que Z_t seja correlacionado com as variáveis independentes) será menor do que 1. A equação pode ser usada para prever Z_t .

Subtraindo-se os valores calculados (preditos) de Z_t dos valores atuais, nós podemos calcular uma série residual, a_t , que representa movimentos não

explicados em Z_t , isto é, o ruído puro. Uma aplicação da análise de séries temporais é construir um modelo ARIMA para a série residual a_t da regressão. Então, pode-se substituir o erro implícito na equação de regressão original pelo modelo ARIMA. Ao usar-se a equação para prever Z_t , pode-se também fazer uma previsão do erro e_t , usando-se o modelo ARIMA. O modelo ARIMA fornece algumas informações sobre os valores futuros de e_t ; isto é, ajuda a explicar a variância inexplicada na equação de regressão. A combinação do modelo de regressão-séries-temporais é

$$Z_t = v_0 + v_1 x_{1Z} + v_2 x_{2Z} + \mathbf{f}^{-1}(B)\mathbf{q}(B)\mathbf{h}_t \quad (2.39)$$

onde \mathbf{h}_t é o termo do erro normalmente distribuído no qual pode ter variância diferente do termo e_t . É provável que este modelo faça previsões melhores que o modelo de regressão encontrado pela equação (2.38), ou por um modelo de série temporal, desde que inclua uma explicação estrutural por parte da variância de Z_t , que não pode ser explicado estruturalmente (Pindyck e Rubinfeld, 1991).

A equação (2.39) é um exemplo do que às vezes chama-se modelo de função de transferência ou, alternativamente, um modelo Autorregressivo-Média-Móveis-Multiplicativo (modelo MARMA). Um modelo de função de transferência relaciona a variável dependente com seus valores anteriores, os valores atuais e valores anteriores de uma ou mais variáveis independentes, e um erro que é parcialmente explicado por um modelo de séries-temporais. Assim, a forma generalizada para um modelo de função de transferência univariado (só uma variável independente) pode ser escrita como

$$Z_t = v^{-1}(B)\mathbf{v}(B)x_t + \mathbf{f}^{-1}(B)\mathbf{q}(B)\mathbf{h}_t \quad (2.40)$$

A técnica de modelagem de uma função de transferência envolve a análise das funções de autocorrelação (FAC) e autocorrelação parcial (FACP) para a variável independente x_t , como também a variável dependente Z_t , em um esforço para especificar o *lag* polinômios $v(B)$, $w(B)$, $f(B)$, e $q(B)$ (Box e Jenkins, 1976). Um problema com a técnica, porém, é que a especificação da parte estrutural do modelo (os polinômios $v(B)$ e $w(B)$) só é feito mecanicamente.

A combinação do uso da análise de regressão com modelos de séries-temporais do termo do erro em alguns casos provêm a melhor previsão. Demonstração e exemplos desta técnica podem ser encontrados em Pindyck e Rubinfeld (1991).

2.5.1 Modelos de função de transferência de Box & Jenkins com entrada simples

Conforme Wei (1990), a formulação da relação linear que relaciona as séries discretas Z_t e X_t :

$$Z_t = \sum_{j=0}^{\infty} v_j X_{t-j} \quad (2.41)$$

onde $Z_t = (v_0 + v_1 B + v_2 B^2 + \dots) \cdot X_t$

ou $Z_t = v(B) \cdot X_t$

A equação (2.41) denomina-se Filtro Linear ou Modelo de Função de Transferência Linear. O polinômio de grau infinito em B , $v(B)$ é conhecido como Função de Transparência de Filtro. Os pesos v_0, v_1, v_2, \dots são chamados respostas impulsiais ou função de resposta de impulso do sistema.

O sistema diz-se estável quando a série $v(B)$ converge para $|B| \leq 1$. A condição de estabilidade implica que mudanças finitas na série de entrada resultam em mudanças na série de saída. Às vezes, o efeito de uma mudança na série de entrada não se manifesta imediatamente sobre a série de saída, mas após “b” instantes de tempo:

$$Z_t = v(B) \cdot X_{t-b} \quad (2.42)$$

$$\text{onde: } X_{t-b} = B^b \cdot X_t \quad (2.43)$$

e o parâmetro inteiro “b” é denominado defasagem.

O modelo pode conter ainda, um ruído, η_t , independente de X_t :

$$Z_t = v(B) \cdot X_{t-b} + \eta_t. \quad (2.44)$$

O ruído η_t , em geral, não é ruído branco, e pode ser representado por um modelo Autorregressivo-Integrado-Média-Móveis (ARIMA).

O modelo (2.43) que relaciona Z_t e X_t , conforme Wei (1990), contém um número infinito de pesos (v_0, v_1, \dots), portanto, não é parcimonioso. Assim, se $v(B)$ for uma função racional, pode-se escrevê-la:

$$v(B) = \frac{\omega(B)}{\delta(B)} \quad (2.45)$$

onde: $\omega(B) = \omega_0 - \omega_1 B - \omega_2 B^2 - \dots - \omega_s B^s$ e $\delta(B) = 1 - \delta_1 B - \delta_2 B^2 - \dots - \delta_r B^r$

são polinômios de graus “s” e “r”, respectivamente.

Função de correlação cruzada - A ferramenta utilizada na análise de dados para a identificação do modelo de função de transferência é a função de correlação cruzada entre a entrada e a saída. A dependência normalizada entre as séries X_t e Z_t será medida pela função de correlação cruzada, denotada por $\rho_{xy}(k)$, dada por :

$$\rho_{xy}(k) = \frac{\text{Cov}[X_t, Y_{t+k}]}{[\sigma_x^2 \sigma_y^2]^{1/2}} = \frac{\gamma_{xy}(k)}{[\gamma_{xx}(0) \gamma_{yy}(0)]^{1/2}} \quad (2.46)$$

com $k \in \mathbb{Z}$, sendo σ_x e σ_y os desvios padrões das séries X_t e Z_t , respectivamente (Brockwell e Davis, 1987).

2.5.2 Teste de causalidade

O conceito de causalidade é fundamental na construção de modelos de função de transferência. Poucos se questionam da verdadeira direção de causalidade entre as variáveis ao desenvolver um modelo. No entanto, ao colocarem uma variável como função de outras, ditas independentes, estão fazendo uma forte hipótese com relação à causalidade entre aquela e estas. Quando existem relações causais unidirecionais entre duas ou mais séries temporais, pode-se construir modelos de função de transferência que unem estas variáveis. Estes modelos entre duas variáveis X e Z só têm sentido se X causa Z . Além disso, é

necessário que o inverso não ocorra, isto é, se Z também causa X tem-se uma relação de “feedback” e os modelos de função de transferência não serão mais adequados.

Pode-se definir causalidade de maneira formal e informal. Informalmente, Granger e Morris (1976) definem que a variável X causa a variável Z se o valor atual de Y é melhor predito utilizando-se valores passados de X do que não os utilizando, isto é, mantendo-se iguais todas as outras informações.

Um dos passos na construção de testes para a direção da causalidade é relacionar a definição de Granger com a estrutura dos modelos paramétricos relacionados às séries temporais.

Teste direto de Granger - O teste implícito na definição de Granger (1969), consiste em utilizar um modelo bivariado AR(M) de ordem M suficientemente grande para aproximar a representação vetorial AR(∞), que é,

$$X_t = \sum_{j=1}^M \mathbf{p}_{11j} X_{t-j} + \sum_{j=1}^M \mathbf{p}_{12j} Z_{t-j} + \mathbf{a}_t, \quad (2.47)$$

$$Z_t = \sum_{j=1}^M \mathbf{p}_{21j} X_{t-j} + \sum_{j=1}^M \mathbf{p}_{22j} Z_{t-j} + \mathbf{h}_t, \quad (2.48)$$

Por exemplo, a hipótese nula de que Z não causa X é equivalente a hipótese $\mathbf{p}_{12j} = 0 \quad j=1, \dots, M$ na regressão (2.47). Esta hipótese é testada usando-se o teste F.

Este procedimento tem o defeito de que um grande número de graus de liberdade (observações) são perdidas ao estimar os parâmetros do modelo (2.47) ou (2.48) que é usado para aproximar a estrutura do correspondente modelo AR(∞). Uma alternativa é tentar obter um modelo parcimonioso da forma mista ARMA(p, q) vetorial. A hipótese nula, Z não causa X , pode ser testada através do teste da razão de verossimilhança da hipótese:

$$\mathbf{f}_{12}(L) = 0 = \mathbf{q}_{12}(L), \quad (2.49)$$

na representação ARMA vetorial
$$\begin{bmatrix} X_t \\ Z_t \end{bmatrix} - \sum_{j=1}^p \mathbf{f}_j \begin{bmatrix} X_{t-j} \\ Z_{t-j} \end{bmatrix} = \begin{bmatrix} \mathbf{a}_t \\ \mathbf{h}_t \end{bmatrix} - \sum_{j=1}^q \mathbf{q}_j \begin{bmatrix} \mathbf{a}_t \\ \mathbf{h}_t \end{bmatrix} \quad (2.50)$$

que toma forma:
$$LR : n \log \frac{\left| \hat{\Sigma}_r \right|}{\left| \hat{\Sigma} \right|} \quad (2.51)$$

e é assintoticamente distribuída como uma χ^2 com $p+q$ graus de liberdade. Em (2.51) $\hat{\Sigma}$ e $\hat{\Sigma}_r$ são as estimativas das matrizes de variância e covariância dos resíduos do modelo completo (2.50) e restrito (2.49), respectivamente.

O teste de causalidade instantânea é obtido testando-se a matriz da variância dos resíduos e (Σ) é diagonal.

2.6 COMENTÁRIOS GERAIS DO CAPÍTULO

Neste capítulo, procurou-se apresentar o suporte teórico descrito por vários autores da área de Controle Estatístico de Qualidade, de Modelos de Séries Temporais e de Modelos de Regressão de Poisson. Foram apresentados os procedimentos para se usar as técnicas de controle estatístico de qualidade em observações correlacionadas. Foram abordados os principais modelos de séries temporais para a monitoração do processo. E foram apresentados os modelos de regressão de Poisson para a aplicação no controle estatístico do processo. Os modelos serão utilizados para a identificação das variáveis geradoras de instabilidade do sistema.

III – APRESENTAÇÃO E ANÁLISE DOS DADOS

Neste capítulo, faz-se a apresentação e aplicação da metodologia proposta a um conjunto de dados reais. Na seção 3.1, descreve-se a coleta dos dados analisados. Na seção 3.2, realiza-se um estudo para verificar a autocorrelação das variáveis contínuas e das variáveis discretas, no período de julho a setembro de 2001. Na seção 3.3, realiza-se um estudo para verificar a autocorrelação das variáveis contínuas e das variáveis discretas, no período de fevereiro a abril de 2002. Tanto na seção 3.2, quanto na seção 3.3 é feito um ajuste utilizando-se dos resíduos das equações encontradas. Na seção 3.4, é feito o comentário do capítulo.

3.1 COLETA DOS DADOS

A Indústria Têxtil Oeste Ltda. está localizada no município de Mondai, no oeste de Santa Catarina. Ela adquiriu sua primeira Extrusora para fabricação do fio que tece os sacos de polipropileno em 1987. Na época, a produção de sacos estava sujeita a fornecedores, uma vez que a empresa não fabricava a matéria prima.

Eram consumidos 45 mil quilos de fio por mês, vindos de Curitiba, os quais resultavam em 550 mil sacos/mês. “A aquisição da máquina nos permitiu aceitar pedidos maiores e produzirmos o nosso próprio fio, ao invés de depender de fornecedores”, diz Henrique Deiss, Diretor da empresa.

Passados treze anos, a indústria adquiriu mais duas extrusoras. Hoje, a Têxtil Oeste possui 48 teares que consomem em média 3374,99 toneladas de grãos por mês e, em consequência disso, produz 26,17 milhões de sacos/mês. Sendo uma empresa de grande porte, a maior do seu município e do ramo em Santa Catarina. Os seus clientes estão distribuídos por todo o estado de Santa Catarina, do Rio Grande do Sul, do Mato Grosso, do Paraná e de São Paulo.

As séries utilizadas neste estudo são séries diárias coletadas em dois períodos distintos de tempo:

- 1º período – julho a setembro de 2001

- 2º período – fevereiro a abril de 2002

As séries estudadas foram:

a) Número de ocorrências, por tipo de não-conformidade, em bobinas no setor de tecelagem – *dados discretos* – Diariamente, é verificado o número de não conformidades encontradas no tecido de polipropileno os quais são utilizadas como embalagens de diversos produtos e alimentos (como ração, adubo orgânico, etc.). Salienta-se que a fita de polipropileno produzida leva em torno de um a dois dias para ser transformado em tecido. Devido a esta defasagem, a série utilizada no primeiro período inicia em agosto de 2001 até setembro de 2001 (variável dependente). A série utilizada no segundo período, devido à defasagem, inicia em março de 2002 até maio de 2002 (variável dependente). As causas de não conformidades são passamento errado, falta fita urdume e trama que cai.

b) Controle do DENIER¹ da fita – *dados contínuos* – Após a confecção da fita de polipropileno, é feito o controle da resistência da fita, se estiver dentro das especificações é feito o controle da gramatura (DENIER). Verificam-se as especificações da gramatura da fita de polipropileno, se estiver dentro das especificações, a fita vai para a tecelagem. Devido à defasagem de um a dois dias para o fio chegar no setor de tecelagem, a série utilizada no primeiro período inicia em julho de 2001 até setembro de 2001. A série utilizada no segundo período inicia em março de 2002 até maio de 2002. Estas séries são explicativas, pois a variabilidade no fio ocasiona variação no tecido.

As observações servem para a identificação e estimação dos modelos e para a análise da capacidade preditiva do modelo.

O sistema de produção da fábrica desde o recebimento do grão de polipropileno até a confecção do tecido está classificado em três partes, conforme fluxograma da figura 5, página 47. A primeira parte, então, é o recebimento do grão, que depois de colocado no almoxarifado é feita uma amostragem, que é colocada no *plastômetro*², onde é realizado o teste para verificar o índice de fluidez. São coletadas 4 amostras de 5 gramas, ou seja, quatro medidas que são pesadas e verificadas as especificações (①). Depois, são liberadas e vão para a segunda fase do processo, que é a confecção da fita de polipropileno (ráfia) pelas extrusoras, as

¹ DENIER – é a titulação do peso em 9.000 metros de fita que vem em bobinas

² Plastrômetro - instrumento de ensaio de determinação do índice de fluidez de polipropileno

quais são divididas em três grupos chamados de linha 1, linha 2 e linha 3. A linha 1 é responsável pela confecção da fita urdume (fita de polipropileno usada no sentido vertical no transamento do tecido), e as demais linhas são responsáveis pela confecção da fita trama (fita de polipropileno usada no sentido horizontal no transamento do tecido).

Um conjunto de 13 bobinas destas fitas de polipropileno é considerado para a amostragem, a qual é coletada no início de cada um dos três turnos de trabalho da indústria. Com estas 13 bobinas, são retiradas as amostras de 90m para a determinação da resistência, tração e alongamento do fio, utilizando-se a norma NBr 11912 da ABNT/Abril92 (Anexo D). Passando pelas especificações (②), estas amostras passam pela *aspa motorizada*¹, onde são pesadas na balança de precisão conforme as normas NBr5426 da ABNT/Jan85 (Anexo D). Nesta etapa, verifica-se se a fita atende às especificações da titulação DENIER (③) da linha de produção. Caso não atenda às especificações, a fita é descartada para ser retrabalhada até atingir o valor adequado.

Depois de analisada a não conformidade dos mesmos, a fiação é liberada para os 48 teares, os quais produzem o tecido para a confecção dos sacos de polipropileno. Estes tecidos são medidos e verificados se estão dentro das especificações (④) e, caso não estejam, é eliminada a metragem não conforme e verificadas quais suas causas de não conformidades.

Na figura 5, mostra-se o fluxograma do processo.

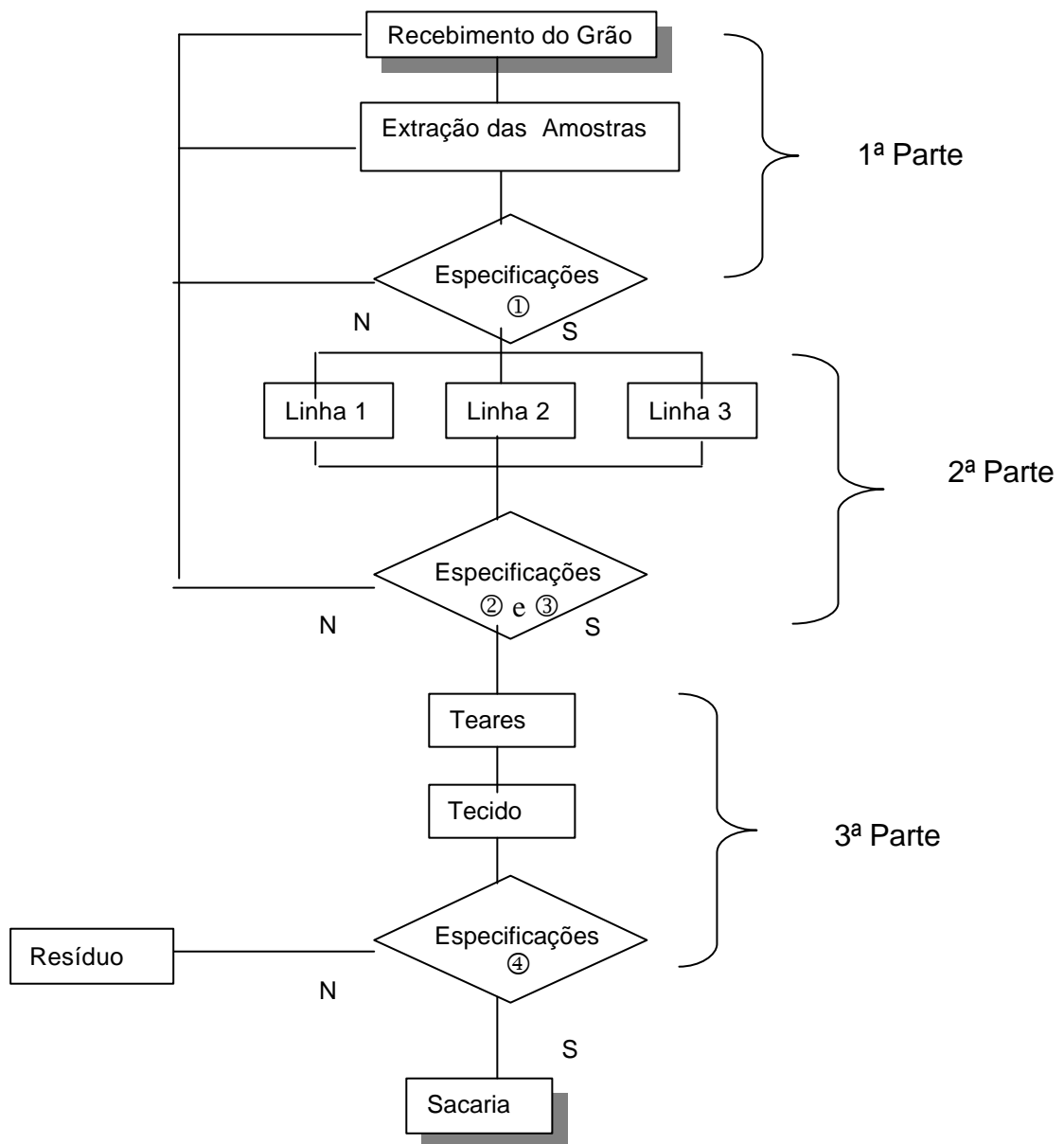


FIGURA 05 - Fluxograma do Processo

A seguir, estão descritos os problemas de maior ocorrência:

- *Passamento errado*: quando há um erro no passamento do fio;
- *Falta fita urdume*: quando o fio termina/ou arrebenta;
- *Trama caiu*: quando o fio se solta e cai fora da bobina.

No processo de confecção da fita de ráfia, há várias situações críticas que devem ser consideradas. Uma dessas situações é a gramatura da fita. Num esforço

para desenvolver um plano de controle para assegurar qualidade da superfície adequada da fita, foi determinado que a gramatura tem um impacto principal na qualidade da fita. Assim, para controlar a qualidade da fita de ráfia, deve-se controlar a gramatura da mesma. Antes de se aplicar os gráficos de controle de Shewhart, três suposições fundamentais devem ser conhecidas:

1. O processo está sob controle;
2. Os dados são normalmente distribuídos;
3. As observações são independentes.

Montgomery (1997) considera que os pontos que estão fora dos limites das especificações são determinados razoavelmente bem pelos gráficos de controle de Shewhart quando a suposição de normalidade é um pouco violada, mas quando as observações não são independentes, os gráficos de controle produzem resultados enganosos. Infelizmente, muitos processos não produzem observações independentes. Alwan (1991) descreve um método para gráficos de controle com dados autocorrelacionados. O método envolve o ajuste de uma curva de séries temporais e depois um gráfico para o controle dos resíduos.

O estudo foi realizado em três etapas:

1ª Etapa – Análise da série gramatura (DENIER) da fita de polipropileno – Foi utilizada a série contínua retirada da segunda parte do processo, na especificação de número 3. Foi feita uma análise exploratória destes dados observados na gramatura da fita; foi ajustado um modelo de série temporal nos mesmos (estimando-se os valores da série) e, depois, foi verificado o controle, utilizando-se as técnicas de Controle Estatístico do Processo (CEP); conforme descrito no capítulo 2. Resumindo, a primeira etapa foi realizada por meio de métodos de previsão que indicarão o conhecimento futuro das variáveis, e depois a utilização por meio dos gráficos de controle.

2ª Etapa – Análise da série não-conforme no setor de tecelagem – Foi utilizada a contagem da série binária (conforme, não conforme), da terceira parte do processo, na especificação de número 4. Foi feita uma análise exploratória destes dados observados no setor de tecelagem, os quais serão ajustados através da análise de

regressão de Poisson. Com os resíduos do modelo, foi verificado o controle, utilizando-se as técnicas do Controle Estatístico de Processos.

3ª Etapa – Ajuste Final – Ao se confirmar a suposição de que a série gramatura da fita de polipropileno é uma variável explicativa da série não-conforme no setor de tecelagem, foi necessário o ajuste final pela combinação dos dois modelos via função de transferência. Este ajuste foi feito modelando-se via função de transferência os resíduos encontrados na série tecidos e para a série gramatura da fita. Como a fita produzida leva em torno de um a dois dias para ser transformado em tecido, o modelo utilizado foi do tipo:

$$Z = \mu + \text{tecido defasado} + \text{gramatura}$$

onde: μ = parâmetro representativo do nível ou média do processo.

Como foi exposto na metodologia, o estudo foi realizado durante dois períodos distintos de tempo, no ano de 2001 e no ano de 2002. No primeiro período em 2001 os dados foram coletados de 31 de julho a 30 de setembro de 2001; e no segundo período os dados foram coletados de 28 de fevereiro a 30 de abril de 2002. A metodologia estudada em 2001 foi reaplicada em 2002 a fim confirmar o estudo.

3.2 ANÁLISE DOS DADOS – ANO 2001

Os dados utilizados para análise nesta seção foram: a série gramatura (DENIER) da fita de polipropileno (variáveis contínuas) e a série de não conformidade no setor de tecelagem (variáveis discretas). Foi feita uma análise exploratória destes dados; foi ajustado um modelo para cada série separadamente, depois foi feita a análise de transferência e verificado o controle, utilizando-se as técnicas de Controle Estatístico do Processo (CEP).

3.2.1 Análise da série gramatura da fita de polipropileno – ano 2001

A análise segue os passos do fluxograma da figura 5 descrita no item 3.1, página 44. Nas partes 2 e 3 do processo, foi efetuado um estudo que ajuda a verificar onde está a maior instabilidade do processo, para que se possa fazer um controle melhor do sistema. Os dados na tabela 2 são os 61 valores médios diários

da gramatura da fita. O valor do dia 30 de setembro foi desconsiderado para a análise devido à defasagem.

TABELA 2 - Valores diários da gramatura da fita coletados no período de 31 de julho a 30 de setembro de 2001

Data	Valores
31-Julho	685,5
Agosto	
1-10/ago	700,90 668,44 679,59 674,05 678,51 671,03 674,13 680,87 679,18 672,54
11-20/ago	674,54 669,28 672,69 676,77 657,31 674,87 685,82 682,33 683,90 680,08
21-30/ago	673,00 662,04 668,41 645,73 645,26 660,54 650,87 629,23 663,92 663,41
31/ago	653,85
Setembro	
1-10/set	664,59 647,85 652,03 655,9 651,49 687,62 674,62 656,10 663,56 658,18
11-20/set	661,72 648,59 662,33 655,38 659,87 662,56 652,49 648,31 664,56 666,81
21-30/set	654,62 665,18 666,69 656,72 666,31 658,74 668,28 654,62 648,15 653,59

Suspeita-se que os dados diários da gramatura da fita não sejam independentes, e o resultado do gráfico desses valores, como mostra a figura 6, apóia esta convicção. O problema é implementar o controle estatístico para um processo que tem autocorrelação (isto é, que as observações não são independentes).

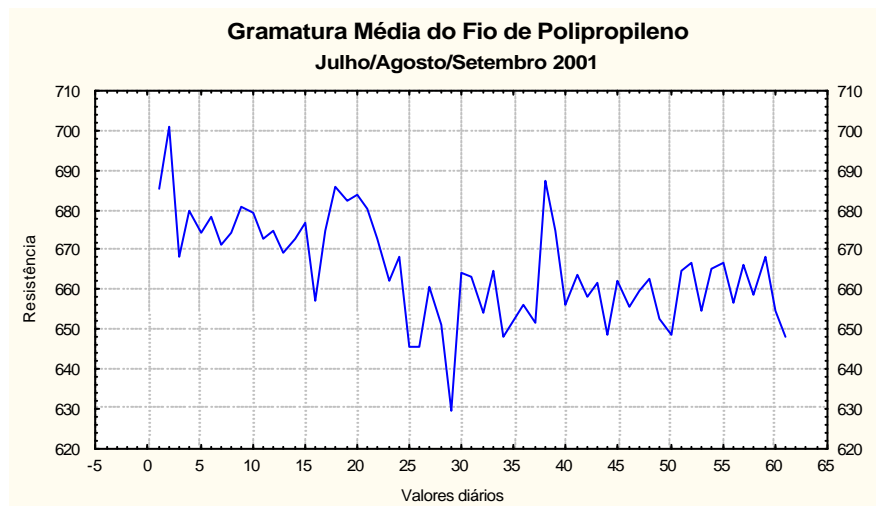


FIGURA 06 – Leitura diária dos dados

A figura 6 mostra a grande variabilidade dos dados, pressupondo que a série não seja estacionária.

Cálculo da autocorrelação

Para confirmar a autocorrelação, cálculos foram executados para confirmar esta suspeita.

A tabela E1, no Anexo E, contém os cálculos requeridos para calcular o coeficiente de autocorrelação para $k=1$ e $k=2$. O coeficiente de autocorrelação para $k=1$ é a soma do produto da coluna 1 e 3 dividida pela soma da coluna 2:

$$r_1 = \frac{4805,8832}{10012,4861} = 0,48$$

O coeficiente autocorrelação para $k=2$ é a soma do produto da coluna 1 por 4 dividido pela soma da coluna 2:

$$r_2 = \frac{3586,4570}{10012,4861} = 0,36$$

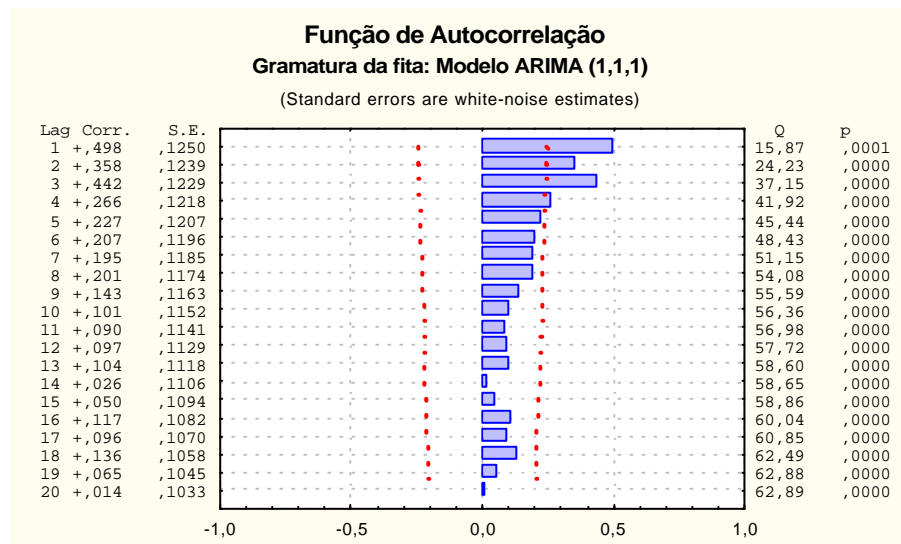


FIGURA 07 – Coeficientes da função de autocorrelação

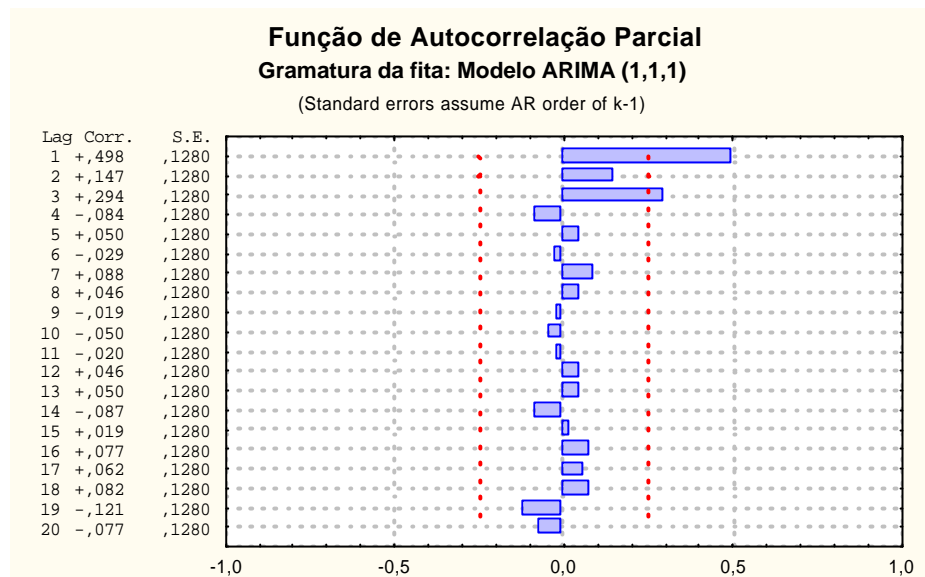


FIGURA 08 – Coeficientes da função de autocorrelação parcial

A figura 7 mostra o coeficiente de autocorrelação e os dois erros padrões para os coeficientes com lags 1 a 4. Como se suspeitava, os dados são altamente correlacionados. Os coeficientes de autocorrelação para os lags 1 a 4 excedem os dois erros padrões. A ACF da série não apresenta nenhuma queda para zero, ratificando, assim, a suposição inicial da inexistência de componente sazonal e da não estacionariedade tendencial da série, quando da análise da figura 6 (dados originais). Antes de se usar os gráficos de controle, deve-se deixar a série estacionária, logo, estes dados devem ser diferenciados para assegurar a estacionariedade e a independência de cada observação.

Removendo a autocorrelação

Para encontrar um conjunto de dados independentes, normalmente distribuídos, Montgomery (1997) recomenda modelar a estrutura da série e desenvolver os gráficos de controle dos resíduos.

Quatro pontos foram usados por causa dos altos coeficientes de autocorrelação dos lags 1 a 4. Para determinar os parâmetros deste modelo, a metodologia de Box e Jenkins foi utilizada. O melhor modelo encontrado foi ARIMA(1,1,1), com o critério de validação MAPE igual a 0,72%.

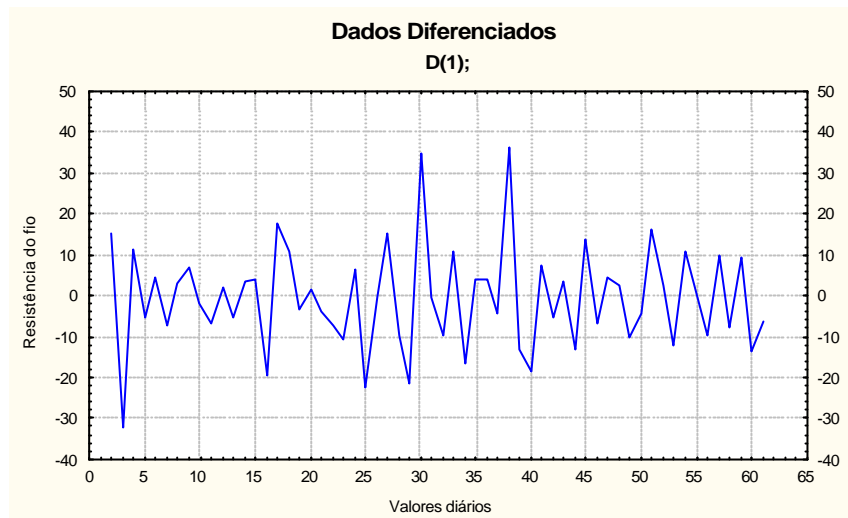


FIGURA 09 – Resíduos da gramatura da fita de rafia

Para cada período de tempo, a diferença do valor atual e do valor previsto da gramatura da fita,

$$\hat{a}_t = Z_t - \hat{Z}_t \quad (3.4)$$

mostra ser uma variável aleatória independente e normalmente distribuída, figura 9.

Aplicando os gráficos de controle

Para validar que a autocorrelação tenha sido removida dos dados, os coeficientes de autocorrelação foram calculados para os dados transformados (diferenciados) definidos pelo modelo ARIMA(1,1,1). Os resultados são mostrados na figura 10.

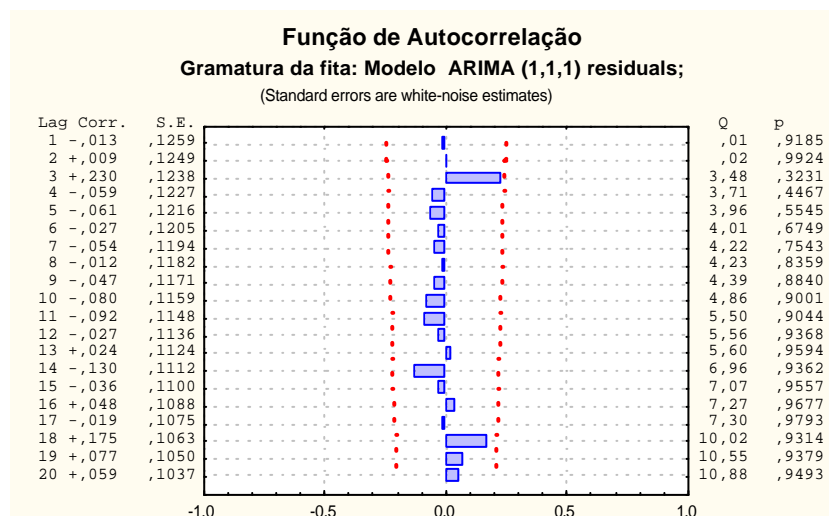


FIGURA 10 – Coeficientes de autocorrelação dos dados transformados

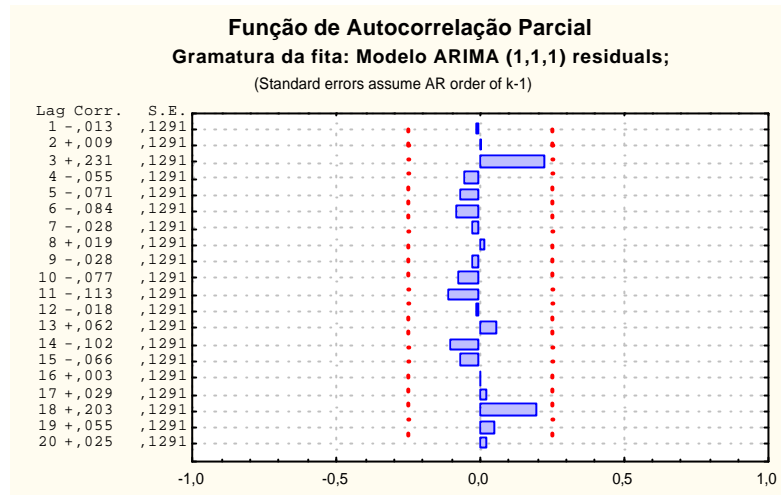


FIGURA 11 – Coeficientes de autocorrelação parcial dos dados transformados

A figura 10 mostra que os dados definidos por \hat{a}_t são independentes de observação para observação. A ACF da figura 10, apresenta uma queda rápida para zero, o que significa que com uma diferença a série tornou-se estacionária. No gráfico da PACF, figura 11, pode-se observar todos os lags dentro dos limites de controle.

O gráfico da probabilidade half normal dos resíduos representados pela figura 12 mostra claramente que os resíduos formam um processo de ruído branco.

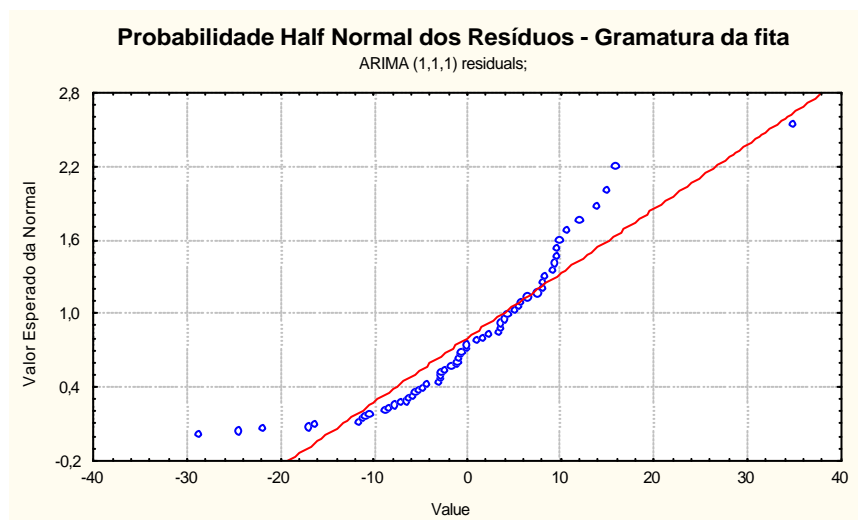


FIGURA 12 - Gráfico representativo da probabilidade half normal dos resíduos

Para utilizar os gráficos de controle de Shewhart, estes dados devem também ser normalmente distribuídos e estarem sob controle. A figura 13 mostra o histograma que indica a aproximação da normalidade dos dados.

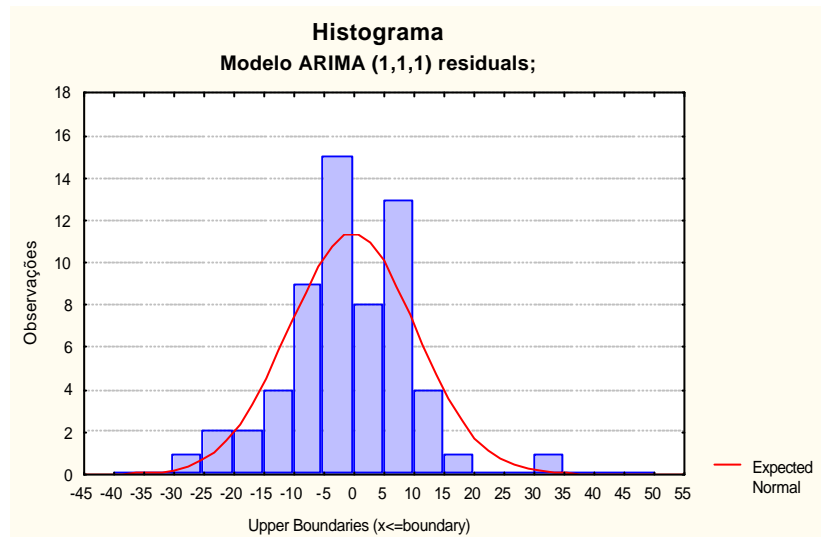


FIGURA 13 - Histograma dos dados transformados

Para a verificação da normalidade, o seguinte teste de qui-quadrado foi executado:

TABELA 3 - Teste do qui-quadrado

Limite Inf	Limite Sup	Obs	Esp	(Obs-Esp) ² /Esp
642,5992	682,7725	658,7436	661,6859	0,0131
639,4955	681,5321	668,2821	660,5138	0,0914
637,5514	682,0950	654,6154	659,8232	0,0411
635,8035	682,5960	648,1538	659,1998	0,1851
634,1248	683,0466	653,5897	658,5857	0,0379
				0,3685

Para a distribuição normal, o número de graus de liberdade é igual ao número de elementos menos 3. Para dois graus de liberdade $c_{0,05}^2 = 5,991$. Como o valor calculado de qui-quadrado foi igual a 0,3685, e é menor do que o valor crítico, os dados apresentam-se numa distribuição normal.

Agora, pode-se verificar o comportamento do processo produtivo. As figuras 14 e 15 mostram as condições de controle para as observações. A figura 14 mostra os gráficos da média (\bar{X}) e da amplitude (R) para os dados originais.

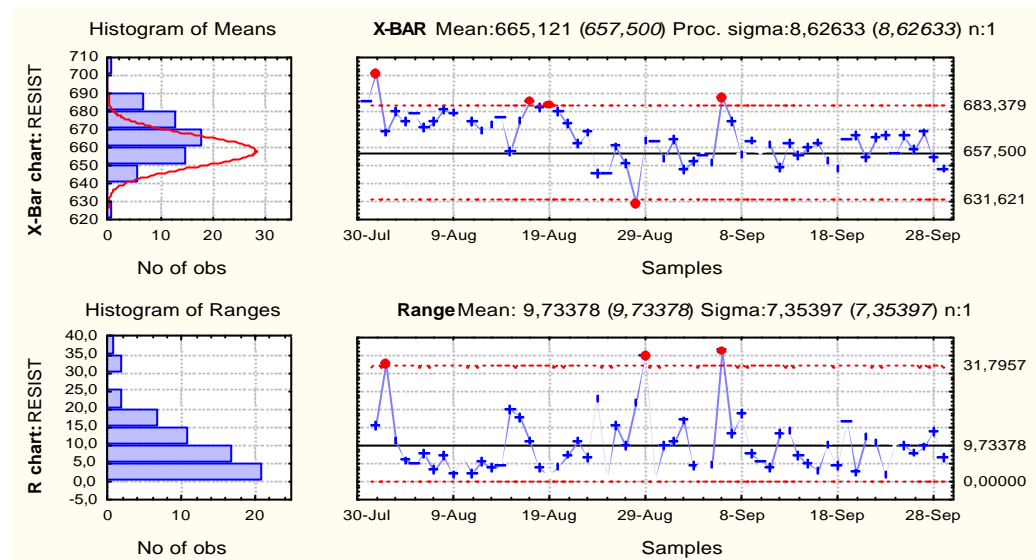


FIGURA 14 - Gráficos da \bar{X} e da R para os dados observados.

Através da figura 14, podemos notar a seqüência de observações e os limites dos gráficos tradicionais de Shewhart, onde vários valores caíram fora dos limites de controle, indicando que o processo está aparentemente fora de controle. Como podemos observar, os dados também aparentam exibir um comportamento sistemático. De fato, antes da transformação (diferenciação) dos dados, os mesmos encontravam-se realmente correlacionados, agora se pode ajustá-los por um processo autoregressivo-integrado-médias-móveis (ou ARIMA) (Wardell, Moskowitz e Plante, 1994).

A figura 15 mostra os gráficos da média (\bar{X}) e da amplitude (R) para os dados transformados (diferenciados). Investigações revelaram que o sistema tinha se dissipado durante esse período de tempo. Após o ajuste dos dados, houve uma ação para corrigir o problema. O problema estava nas primeiras observações, as quais foram ignoradas, a condição de normalidade foi verificada e os gráficos de controle \bar{X} e R foram refeitos.

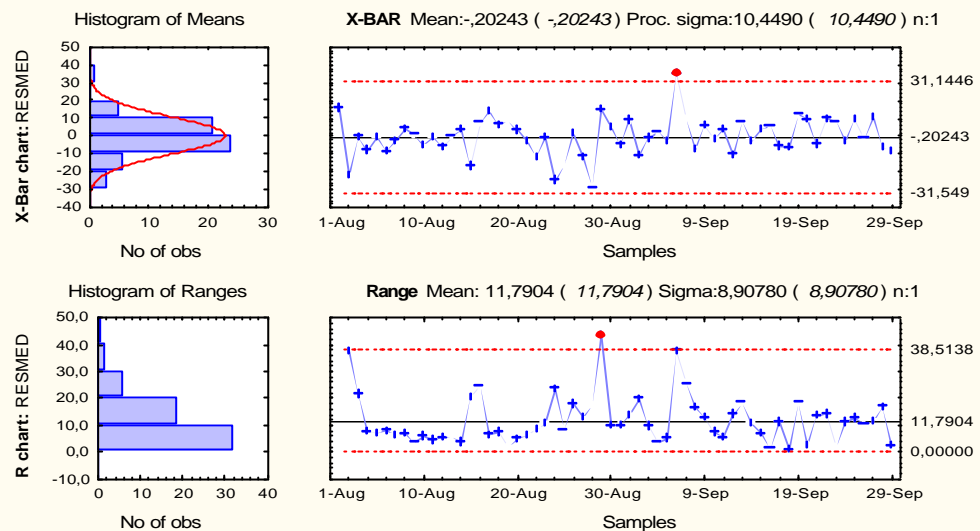


FIGURA 15 - Gráficos da \bar{X} e da R para os dados transformados.

Através da figura 15, podemos observar que os gráficos de controle para os mesmos dados indicam que os valores residuais estão praticamente dentro dos limites de controle para a média. Segundo Wardell, Moskowitz e Plante (1994) isto é inteiramente possível em gráficos de controle tradicionais; os pontos estão fora dos limites por causa da sistemática ou das causas comuns e não por causa de ocorrência de causas especiais.

3.2.2 Análise da série não-conforme no setor de tecelagem – ano 2001

O objetivo desta etapa é aplicar as técnicas descritas na seção 2.4, página 28. A análise segue os passos do fluxograma da figura 5, descrita no item 3.1, página 44. Os dados da tabela 4 são os valores diários da verificação da conformidade no setor de tecelagem da fita de polipropileno.

TABELA 4 - Valores diários coletados do número de itens não conformes no setor de tecelagem no período de 01 de agosto a 31 de setembro de 2001

Data	Valores									
Agosto										
1 a 10/ago	0	1	1	1	0	1	0	2	2	0
11 a 20/ago	0	2	2	0	2	0	0	1	0	1
21 a 30/ago	5	0	0	3	1	1	1	0	0	2
31/ago	1									
Setembro										
1 a 10/set	0	0	2	0	0	0	1	0	0	3
11 a 20/set	0	1	2	5	0	2	1	1	0	0
21 a 30/set	0	2	1	1	1	0	0	3	1	

Na figura 16, mostra-se a série do número diário de não conformidade no setor da tecelagem, no período de agosto a setembro de 2001, num total de 61 valores. Observa-se que a série aparenta ser aleatória, isto é, não apresenta uma tendência.

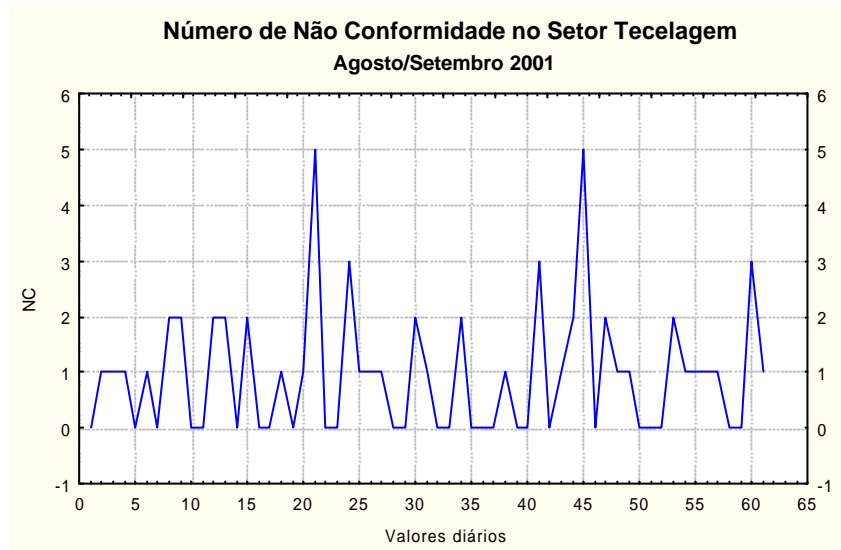


FIGURA 16 – Contagem diária dos dados

Os dados, aparentemente, não apresentam uma tendência sobre o tempo, o que se confirmou mais tarde. Foi aplicado o teste de aderência Kolmogorov-Smirnov, para verificar se os dados podem ser estudados pela distribuição de Poisson.

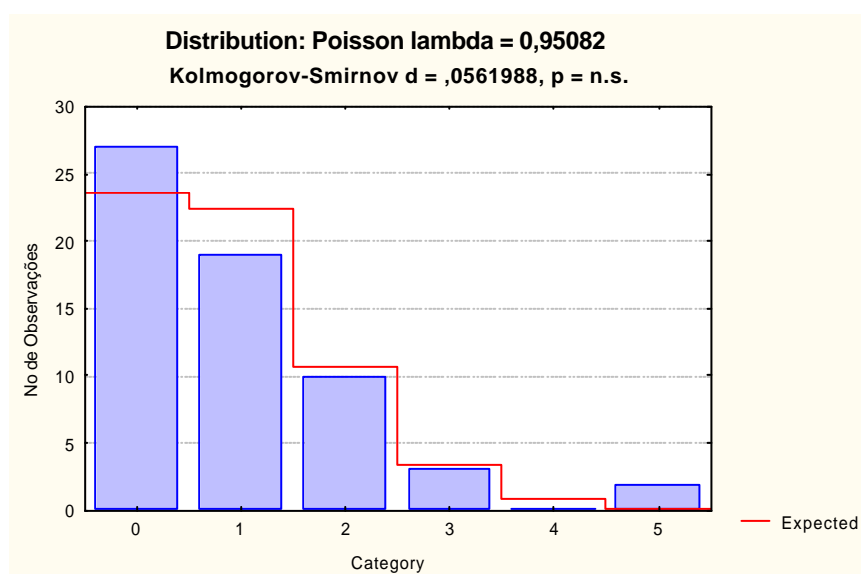


FIGURA 17 – Teste de aderência

Pela figura 17, observar-se através do teste de Kolmogorov-Smirnov que os dados seguem uma distribuição de Poisson, e que possuem uma grande variabilidade. A média de não conformidade é de 0,951, e o desvio padrão 1,161. Foi feito o ajustamento dos dados através do modelo de regressão de Poisson, onde encontrou-se os seguintes resultados:

TABELA 5 – Sumário dos parâmetros do modelo

	coef	Erro Padrão (SE)	coef/SE	valor p
Z	-0.0165	0.1291	-0,13	0.898

O algoritmo de Newton-Raphson convergiu em 5 iterações.

O modelo de regressão de Poisson encontrado foi:

$$Z = -0,0165 + \text{erro}_{(0,1291)}$$

Para verificar a adequação do modelo encontrado, usou-se os testes de Pearson e o desvio G^2 .

A tabela 6 contém informações sobre a avaliação do modelo:

TABELA 6 - Critérios de avaliação do modelo

Critérios	GL - Graus de Liberdade	Valores	Valores/GL
Escala " Deviance" (G^2)	59	81,5738	1,3826
Escala de Pearson (c^2)	59	79,9828	1,3556
Verossimilhança	59	-82,8238	

A Estatística de Desvio (G^2) e de Pearson (c^2), páginas 32 e 33, possuem uma distribuição de c^2 com os mesmos graus de liberdade. A Estatística de Desvio (G^2) e de Pearson (c^2) dividido pelos respectivos graus de liberdade são usados para detectar se há superdispersão ou subdispersão. A média e a variância numa distribuição de Poisson são iguais, isso implica que a Estatística de Pearson (c^2), dividido pelo grau de liberdade, deve ser aproximadamente 1 (um). Valores maiores que um indica uma superdispersão, isto é, a variância real é maior do que a média, valores menores do que um, indicam subdispersão, a variância real é menor do que a média. Evidências de superdispersão ou subdispersão indicam que o modelo de

Poisson não está adequado. Para esses valores, há evidências de uma superdispersão, e a média é 0,9836 e a variância é 1,3163. Todas (100%) as observações têm valores esperados abaixo de 5.0.

O modelo encontrado não é adequado, as possíveis razões seriam:

1. A tendência no tempo é real;
2. Há outras variáveis omitidas no preditor linear; ou
3. Há superdispersão.

Com as presentes informações, não se pode distinguir (2) de (3). Mas pode-se observar a tendência. Suponha-se que exista uma tendência sobre o tempo, necessita-se ter um polinômio de grau 4 para capturar este efeito (ver figura 16, página 55). Tem-se que incluir efeitos no tempo, escolhe-se, então, quatro efeitos:

TABELA 7 - Estimativas do modelo

	coef	Erro Padrão (SE)	coef/SE	valor p
Int	-4,858 10^{-2}	2,443 10^{-1}	-0,20	0,842
t	-1,226 10^{-2}	1,835 10^{-2}	-0,67	0,504
t ²	7,564 10^{-4}	1,660 10^{-3}	0,46	0,649
t ³	2,777 10^{-5}	3,186 10^{-5}	0,87	0,383
t ⁴	-1,242 10^{-6}	2,056 10^{-6}	-0,60	0,546

O modelo de regressão de Poisson encontrado, com quatro efeitos na tendência (instante t, t², t³, t⁴) foi:

$$Z = \underbrace{-4,858}_{(2,443 \cdot 10^{-1})} \cdot 10^{-2} - \underbrace{1,226}_{(1,835 \cdot 10^{-2})} \cdot 10^{-2} t + \underbrace{7,564}_{(1,660 \cdot 10^{-3})} \cdot 10^{-4} t^2 + \underbrace{2,777}_{(3,186 \cdot 10^{-5})} \cdot 10^{-5} t^3 - \underbrace{1,242}_{(2,056 \cdot 10^{-6})} \cdot 10^{-6} t^4 + erro$$

Na tabela 7, verifica-se que os valores de p são maiores do que 0,05, então, pode-se afirmar que os coeficientes não são estatisticamente significativos. Para se verificar a adequação do modelo encontrado, após a inclusão dos efeitos no tempo, deve-se analisar o teste de Pearson e o desvio G^2 .

TABELA 8 - Critérios de avaliação do modelo

Critérios	GL - Graus de Liberdade	Valores	Valores/GL
Escala da "Deviance" (G^2)	59	80,36	1,36
Escala de Pearson (χ^2)	59	77,42	1,31
Verossimilhança	59	-60,36	

A tabela 8 indica que o modelo de Poisson não está adequado, pois a Estatística de Pearson (χ^2) dividido pelo grau de liberdade não é aproximadamente 1 (um), sugerindo evidências de superdispersão ou subdispersão.

Examinando os valores encontrados no gráfico da figura 18, observa-se que o comportamento dos dados indica possível inadequação do modelo, apesar de uma melhora com a inclusão de efeitos no tempo. Como as escalas das estimativas dos GLM não são sempre independentes, um pouco mais de cautela deve-se ter na interpretação do gráfico da Half-Normal (Lewis; Montgomery; Myers, 2001).

O grande valor residual do Desvio (G^2) não indica necessariamente a presença de “outliers” (valores discrepantes). O gráfico da probabilidade meio-normal dos desvios residuais, figura 18, também confirma a ausência de “outliers”.

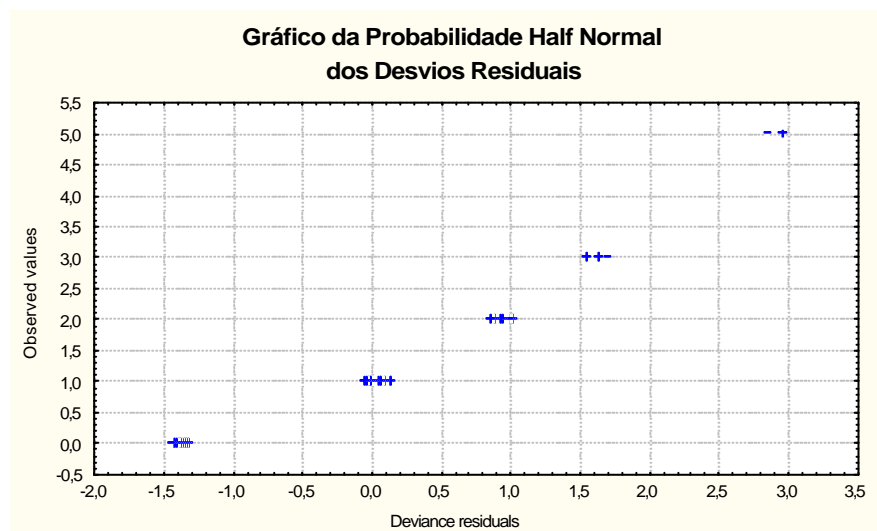


FIGURA 18 – Probabilidade half normal dos desvios residuais

O modelo encontrado não é adequado, sugerindo que a variância da função pode estar errada, isto é, a variância pode simplesmente ser maior do que aquela dada pelo modelo (ver teoria página 34). Um ponto fundamental no processo de ajustamento de um GLM é que não se deve ficar restrito a um único modelo, deve-se fazer várias tentativas (Samohyl, 2001).

Ao analisarmos a superdispersão encontramos os seguintes parâmetros:

TABELA 9 – Sumário dos parâmetros do modelo

	coef	Erro Padrão (SE)	coef/SE	valor p
Z	1,198	0.1486	8,062	0

O algoritmo de Newton-Raphson convergiu em 4 iterações.

O modelo de regressão de Poisson encontrado foi:

$$Z = 1,198 + \text{erro}$$

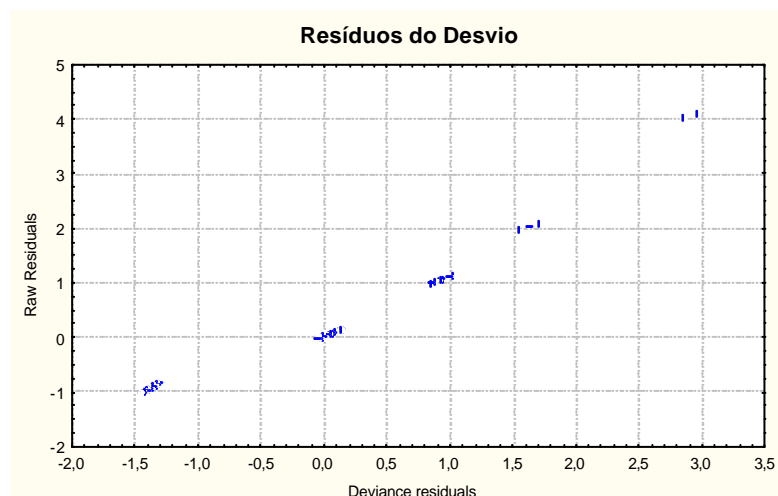
(0,1486)

TABELA 10 - Critérios de avaliação do modelo

Critérios	GL - Graus de Liberdade	Valores	Valores/GL
Escala da "Deviance" (G^2)	59	59,49	1,008
Escala de Pearson (c^2)	59	59,00	1,000
Verossimilhança	59	-64,28	

A Estatística de Pearson (c^2) e o Desvio (G^2) divididos pelos respectivos graus teve-se aproximadamente 1 (um), o que indica que o modelo de Poisson está adequado. Dean e Lawless (1987) sugerem que os resíduos do desvio G^2 sejam preferidos para a avaliação do ajuste do modelo de regressão de Poisson.

Segundo Piegorisch (1998), a análise gráfica e outras operações sobre os resíduos, ajudam a indicar a discrepância do modelo potencial corrente. Box plots dos resíduos do desvio (G^2), dos resíduos de Pearson (c^2) e o gráfico dos resíduos podem indicar o melhor ajuste ou podem identificar efeitos associados com a variável de previsão.

FIGURA 19 – Gráfico dos resíduos do desvio (G^2)

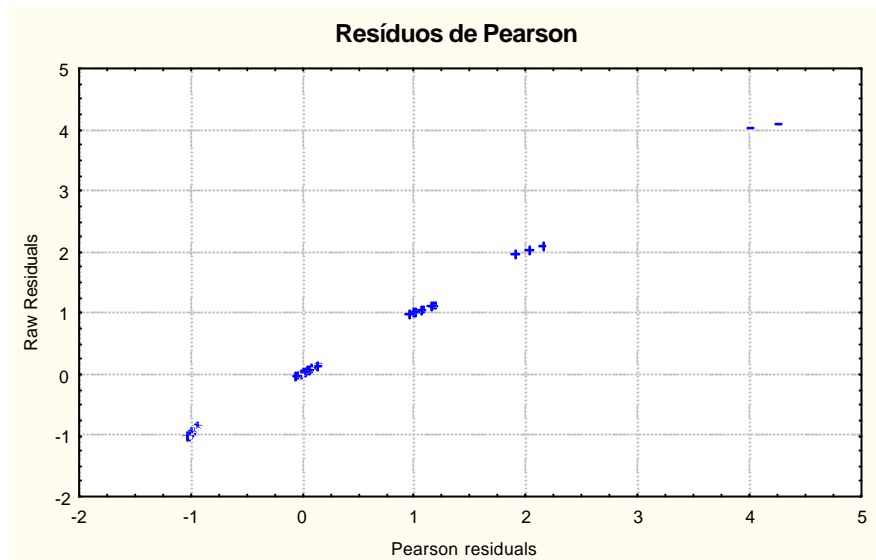


FIGURA 20 – Gráfico dos resíduos de Pearson

Os gráficos das figuras 19 e 20 mostram que ambos os tipos de resíduos são equivalentes para as observações, o que equivale dizer que os resíduos do modelo de regressão de Poisson se apresentam relativamente estáveis (Piegorsch, 1998).

Aplicando os gráficos de controle

Agora, pode-se verificar o comportamento do processo produtivo. As figuras 21 e 22 mostram as condições de controle para as observações. A figura 21 mostra o gráfico (U) para os dados originais.

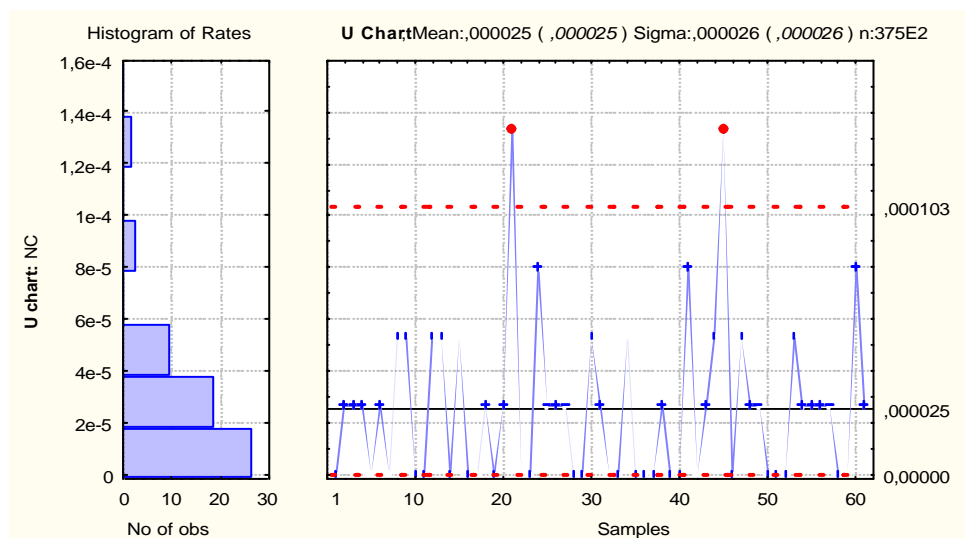


FIGURA 21 – Gráfico U para os dados observados.

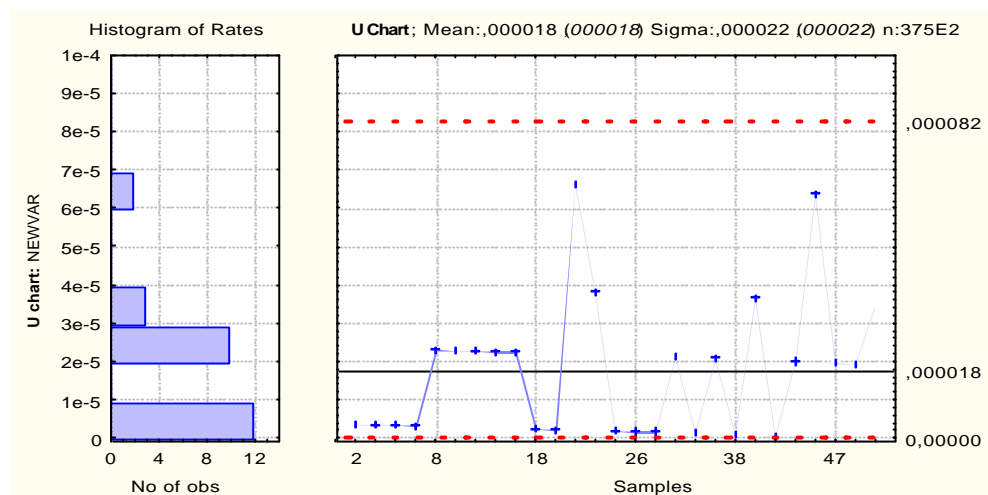


FIGURA 22 – Gráfico U para os dados modelados

Investigações revelam que o sistema não estava ajustado e foi refeito o gráfico de controle U utilizando-se os resíduos encontrados pelo modelo de regressão de Poisson para corrigir o problema. O problema estava nas observações de número 22 e 45 as quais depois de modeladas pela regressão de Poisson foram verificadas dentro dos limites de controle.

Segundo Wardell, Moskowitz e Plante (1994) isto é inteiramente possível em gráficos de controle tradicionais; os pontos estão fora dos limites por causa da sistemática ou das causas comuns e não por causa de ocorrência de causas especiais.

3.2.3 Ajuste final – ano 2001

Nesta etapa, pretende-se unir as duas etapas anteriores (subseção 3.2.1, página 46; e subseção 3.2.2, página 54), de maneira a se conseguir detectar o número possível de não conformidade no resultado do processo. Para isso, aplicou-se o teste de causalidade e ajustou-se um modelo via função de transferência, descrito na seção 2.5, página 34, utilizando-se como variáveis de entrada os resíduos obtidos através da regressão de Poisson e dos modelos Box e Jenkins.

Ao se analisar os resíduos obtidos através da regressão Poisson, observou-se que existe uma forte correlação negativa no lag=2 (figuras 23 e 24) do coeficiente

de autocorrelação e do coeficiente de autocorrelação parcial, indicando que se pode ajustar um modelo autorregressivo.

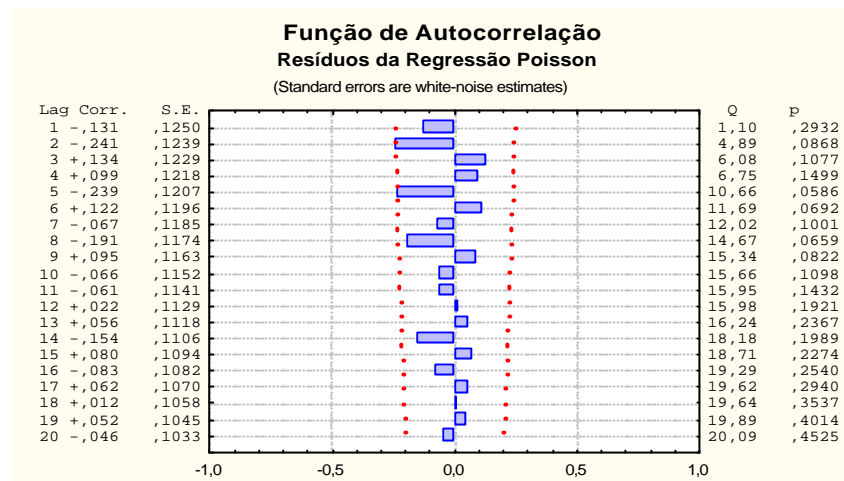


FIGURA 23 – Gráfico da função de autocorrelação

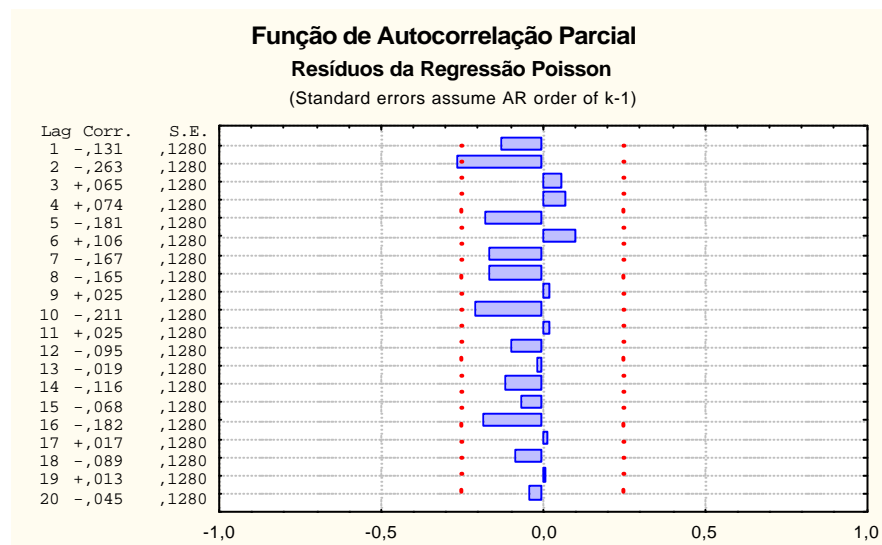


FIGURA 24 – Gráfico da função de autocorrelação parcial

Ao aplicar-se o teste de causalidade descrita na seção 2.5.2, página 38, pode-se afirmar que existe uma relação de causalidade unidirecional instantânea ou nível de significância de 5% para o lag 4 ($p(4)$) no valor de -0,235.

Através da tabela 11, página 62, observou-se que a variável gramatura da fita de polipropileno é uma variável explicativa da variável tecido.

TABELA 11 - Critérios de avaliação do modelo

Fator	Lag	Coef.	Tcal
Média		2,2018	20,27
Autorregressivo	2	-0,35786	-2,74
Gramatura	0	-0,03736	-2,75

O modelo encontrado foi:

$$Z = 2,2018 - 0,35786Z_{t-2} - 0,03736X_t$$

Aplicando os gráficos de controle

Agora, verifica-se o comportamento do processo produtivo. A figura 25 mostra as condições de controle para as observações modeladas através da função de transferência. A figura 26 mostra o gráfico (U) para os dados obtidos com os resíduos da função de transferência.

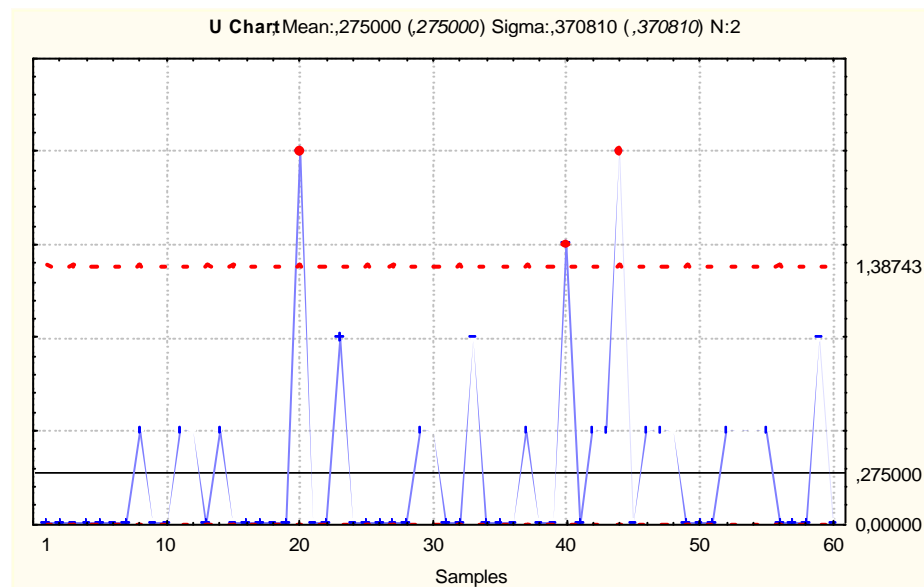


FIGURA 25 – Gráfico (U) para os dados obtidos com os resíduos da função de transferência.

A figura 25 mostra a existência de três pontos fora de controle (amostra 20, 40 e 44) e sugere a retirada destes pontos. Neste gráfico, consegue-se detectar que a amostra 40 está fora dos limites de controle, o que anteriormente não aparecia, pois estava mascarada pela autocorrelação, sendo assim o modelo encontrado através da função de transferência é mais sensível aos dados.

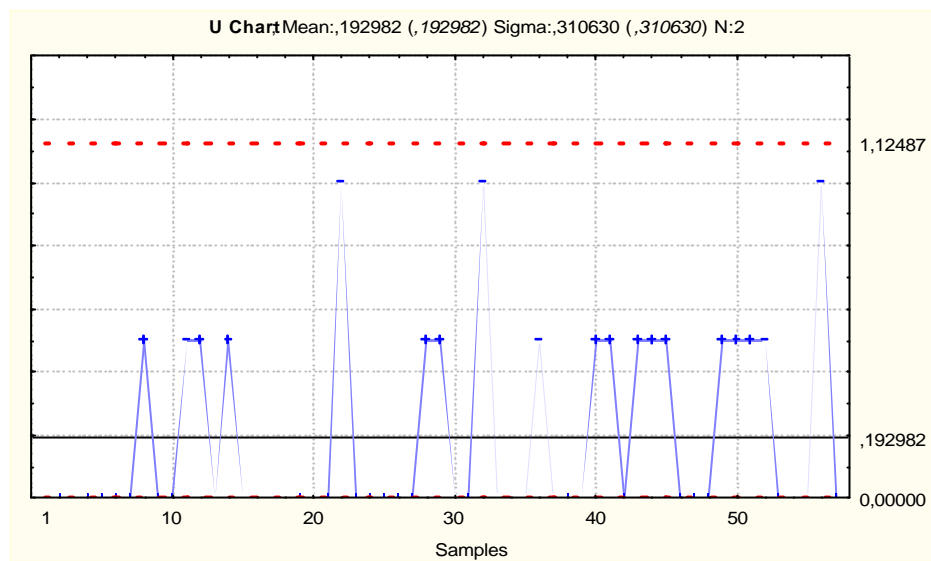


FIGURA 26 – Gráfico (U) para os dados obtidos com os resíduos da função de transferência, sem a amostra 40

Na figura 26, observa-se que, com a retirada dos pontos que estavam fora dos limites controle (amostra 20, 40, 44), o processo ficou sob controle.

3.2.4 Comentários da análise dos dados – ano 2001

Além de se verificar a autocorrelação dos dados, pode-se modelar as variáveis discretas através de modelos de regressão de Poisson e as variáveis contínuas através dos modelos Box e Jenkins e, com os resíduos obtidos, utilizar os modelos de função de transferência para se identificar a existência de causalidade. A técnica proposta, de primeiro retirar a autocorrelação dos dados para depois ajustá-los, mostrou-se adequada. Ao se estudar a autocorrelação dos dados, gerou uma nova perspectiva de aprendizagem sobre o processo produtivo através das informações contidas na estrutura de autocorrelação, dos modelos Box e Jenkins e dos modelos de regressão de Poisson, os quais eram ignorados pelo modelo clássico de monitoramento. A função de transferência empregada posteriormente nos resíduos obtidos permitiu a confirmação da causalidade da gramatura do fio com relação ao tecido produzido. Com isso, houve um crescimento de informações para a correta tomada de decisão e pode-se detectar que houve uma melhora nos pontos de saída de controle.

3.3 ANÁLISE DOS DADOS – ANO 2002

As séries usados para análise nesta seção foram as mesmas do ano de 2001, com seus valores coletados em 2002. Então, as séries são: a série gramatura (DENIER) da fita de polipropileno (variáveis contínuas) e a série de não conformidade no setor de tecelagem (variáveis discretas). Foi feita, também, uma análise exploratória destes dados; foi ajustado um modelo para cada série separadamente, depois foi feito a análise de transferência e verificado o controle, utilizando-se as técnicas de Controle Estatístico do Processo (CEP).

3.3.1 Análise da série gramatura da fita de polipropileno – ano 2002

A análise seguirá também os passos do fluxograma da figura 5 descrita no item 3.1, página 44. Nas partes 2 e 3 do processo, foi reaplicado o mesmo estudo da análise no ano de 2001. Os dados na tabela 12 são os 61 valores médios diários da gramatura da fita. O valor do dia 30 de abril de 2002 foi desconsiderado para a análise devido à defasagem.

TABELA 12 - Valores diários da gramatura da fita coletados no período de 28 de fevereiro a 30 de abril de 2002

Data	Valores
28-Fevereiro	655,57
Março	
1-10/mar	665,02 665,33 672,00 658,33 658,50 663,50 660,00 650,83 663,67 664,58
11-20/mar	665,50 650,00 666,00 664,67 658,08 664,06 666,00 660,00 669,50 662,33
21-30/mar	661,67 675,33 649,92 689,00 670,67 661,00 666,50 667,33 667,50 658,17
31/mar	662,67
Abril	
1-10/abr	657,67 660,00 650,00 664,00 662,87 661,19 669,50 663,33 652,67 656,00
11-20/abr	651,00 653,50 665,00 669,33 655,00 677,93 659,64 662,15 664,30 670,00
21-30/abr	664,00 655,00 662,50 671,48 678,75 668,45 651,50 655,45 659,33 669,00

Suspeita-se que os dados diários da gramatura da fita não são independentes, e o resultado do gráfico desses valores, como mostra a figura 27 para fevereiro/março/abril de 2002, apóia esta convicção. O problema é implementar o controle estatístico para um processo que tenha autocorrelação (isto é, que as observações não são independentes).

Comparando a figura 6, página 50, com a figura 27, página 69, nota-se que houve uma redução na variabilidade da série gramatura da fita de polipropileno.

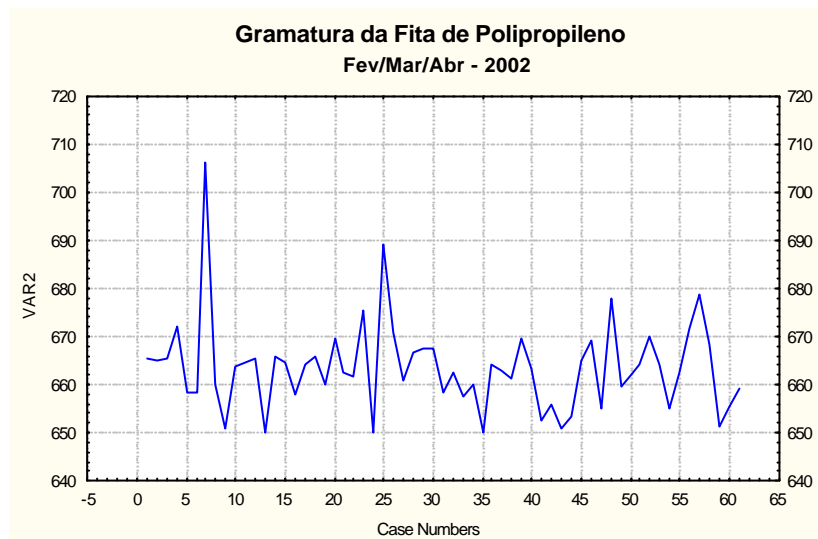


FIGURA 27 – Leitura diária dos dados – ano 2002

Analisando apenas a figura 27, observa-se a grande variabilidade dos dados, e permite verificar a existência de uma série sem tendência, sugerindo a inexistência de componente sazonal significativa (que no decorrer do estabelecimento da ordem do modelo, essa suposição deverá ser analisada, através do comportamento da função de autocorrelação). Pressupõe-se que a série não seja estacionária.

Cálculo da autocorrelação

Por suspeita de autocorrelação, cálculos foram executados para confirmar esta suspeita. A tabela E2, no Anexo E, contém os cálculos requeridos para calcular o coeficiente de autocorrelação para $k=1$ e $k=2$. O coeficiente de autocorrelação para $k=1$ é a soma do produto da coluna 1 e 3 dividida pela soma da coluna 2. O coeficiente autocorrelação para $k=2$ é a soma do produto da coluna 1 por 4 dividido pela soma da coluna 2:

$$r_1 = \frac{-48,4238}{3353,9337} = -0,014$$

O coeficiente autocorrelação para $k=2$ é a soma do produto da coluna 1 por 4 dividido pela soma da coluna 2:

$$r_2 = \frac{34,9061}{3353,9337} = 0,104$$

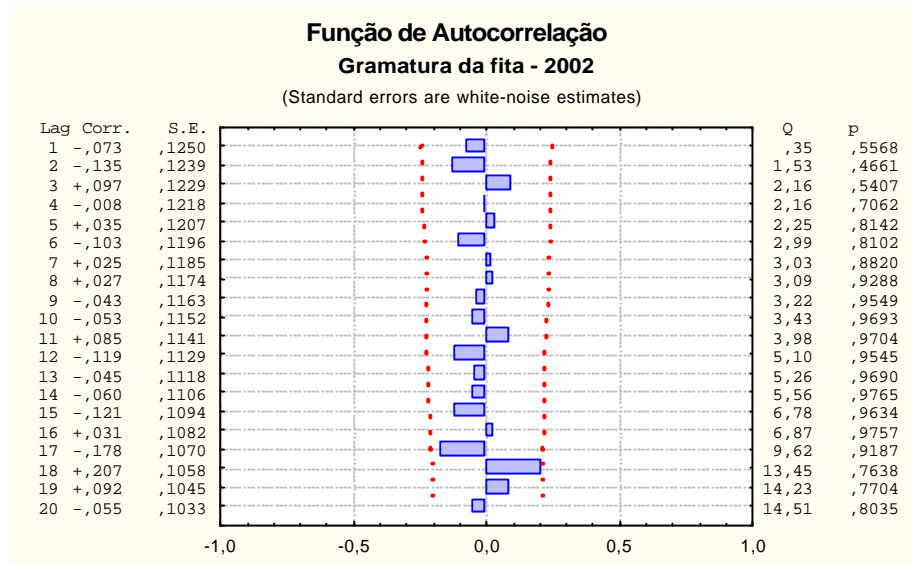


FIGURA 28 – Coeficientes da função de autocorrelação

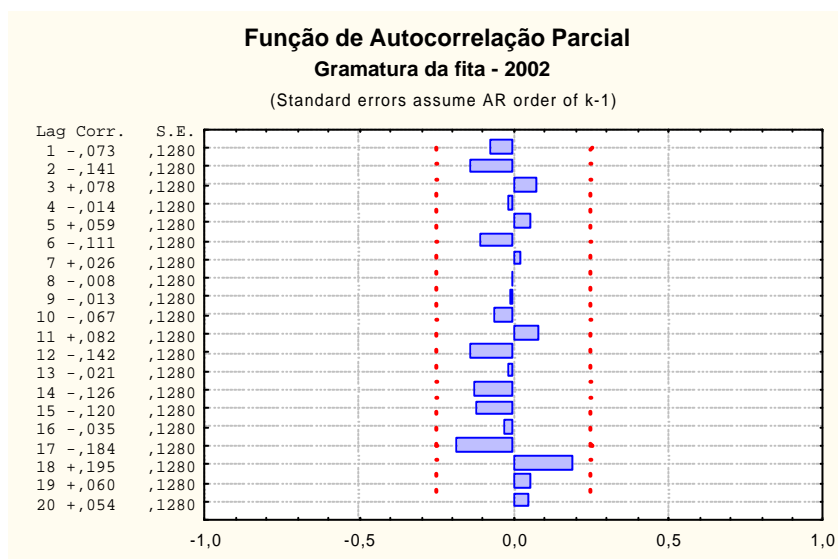


FIGURA 29 – Coeficientes da função de autocorrelação parcial

Pelos valores encontrados para r_1 e r_2 não se identifica correlacionamento, mas a figura 28 mostra a autocorrelação no lag 18. Como se suspeitava, há correlacionamento dos dados. O coeficiente de autocorrelação para os lag 18 excede os dois erros padrão. A ACF da série não apresenta nenhuma queda para zero, ratificando, assim, a suposição inicial da inexistência de componente sazonal e da não estacionariedade tendencial da série, quando da análise da figura 27 (dados

originais). Constata-se também, que os coeficientes de autocorrelação não diminuem rapidamente para zero; pode-se, então, concluir que os dados não são independentes. Antes de se usar os gráficos de controle, sugere-se, então, ajustar um modelo de série temporal adequado às observações individuais, de maneira a deixar a série estacionária.

Removendo a autocorrelação

Os parâmetros do modelo foram obtidos através da metodologia de Box e Jenkins. O melhor modelo encontrado foi ARIMA(1,1,1), com o critério de validação MAPE igual a 1,227%.

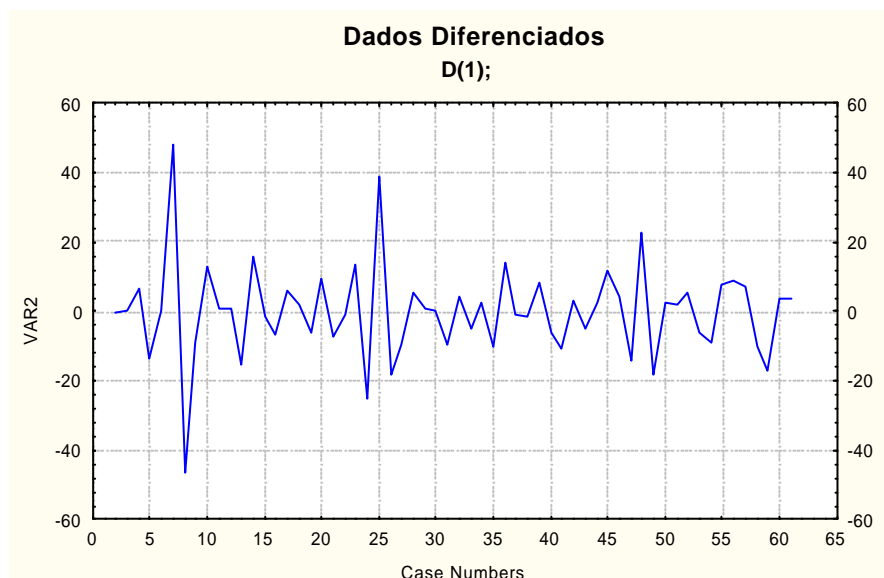


FIGURA 30 – Resíduos da gramatura da fita de rafia

Para cada período de tempo, a diferença do valor atual e do valor previsto da gramatura da fita,

$$\hat{a}_t = Z_t - \hat{Z}_t \quad (3.4)$$

mostra ser uma variável aleatória independente e normalmente distribuída, figura 30.

Aplicando os gráficos de controle

Os coeficientes de autocorrelação foram calculados para os dados transformados definidos pelo modelo ARIMA(1,1,1), a fim de validar que a autocorrelação tenha sido removida. Os resultados são mostrados na figura 31.

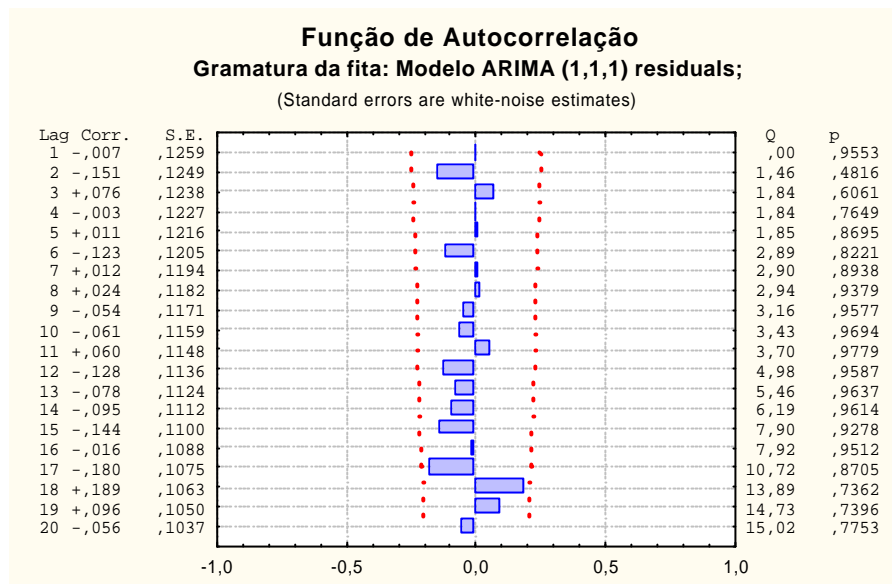


FIGURA 31 – Coeficientes de autocorrelação dos dados transformados

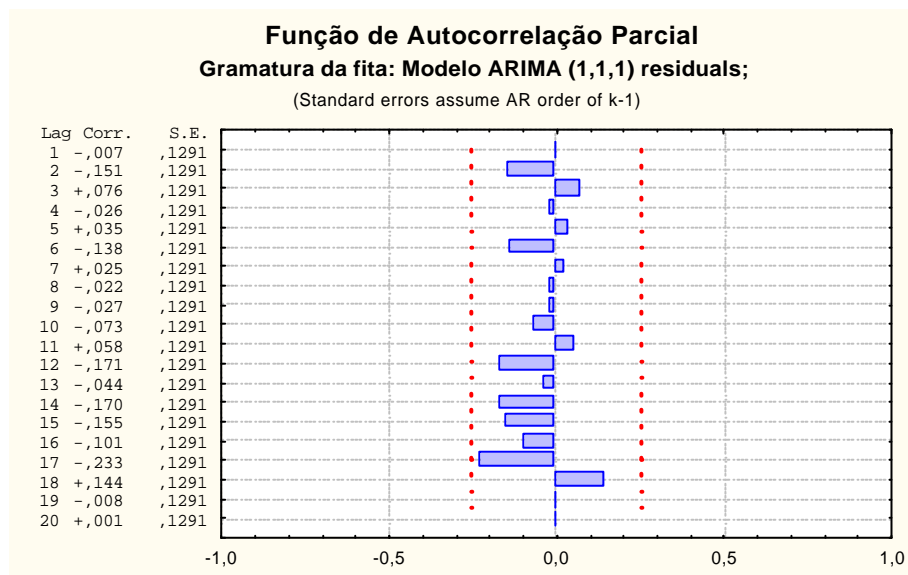


FIGURA 32 – Coeficientes de autocorrelação parcial dos dados transformados

A figura 31 mostra os resultados da análise dos resíduos (\hat{a}_t), onde a função de autocorrelação amostral e o teste Q de *Portmanteau* (ver teoria na seção 2.3.4, página 26) sustentam que os dados são independentes de observação para observação. A PACF da figura 32 apresenta os coeficientes dentro do intervalo de controle.

O gráfico da probabilidade half normal dos resíduos representados pela figura 33 mostra, claramente, que os resíduos formam um processo de ruído branco.

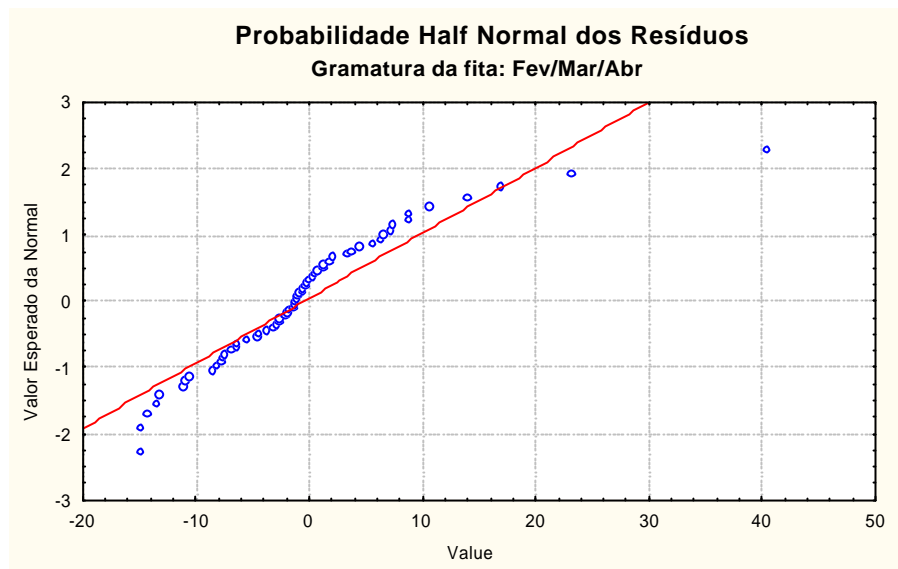


FIGURA 33 - Gráfico representativo da probabilidade half normal dos resíduos

Para utilizar os gráficos de controle de Shewhart, estes dados devem, também, ser normalmente distribuídos e estarem sob controle. A figura 34 mostra o histograma que indica a aproximação da normalidade dos dados.

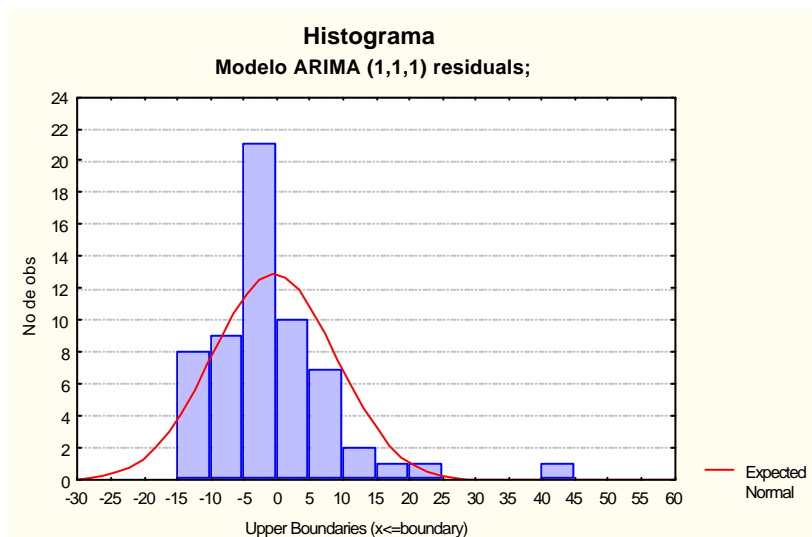


FIGURA 34 - Histograma dos dados transformados

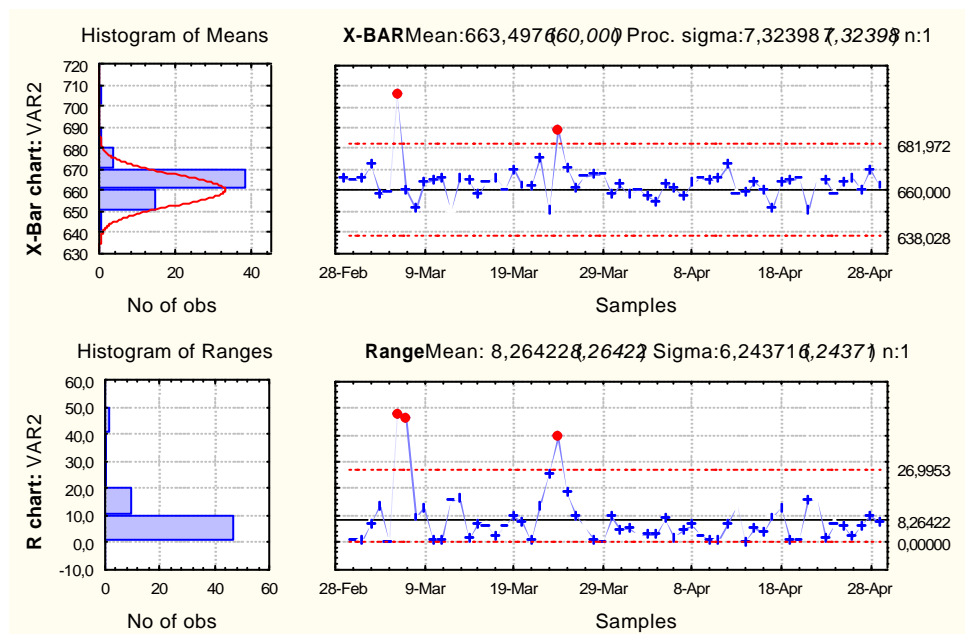
Para a verificação da normalidade, o seguinte teste de Qui-quadrado foi executado:

TABELA 13 - Teste do qui-quadrado

Limite Inf	Limite Sup	Obs	Esp	(Obs-Esp) ² /Esp
646,1137	677,5128	668,4512	661,8133	0,06658
646,9184	678,4644	651,5000	662,6913	0,18900
646,5121	678,5846	655,4565	662,5483	0,07591
646,2321	678,7763	659,3300	662,5042	0,01521
645,9438	678,9573	650,3414	662,4505	0,22135
				0,56804

Para a distribuição normal, o número de graus de liberdade é igual ao número de elementos menos 3. Para dois graus de liberdade $\chi^2_{0,05} = 5,991$. Como o valor calculado de qui-quadrado foi igual a 0,56804, e é menor do que o valor crítico, os dados apresentam-se numa distribuição normal.

Agora, pode-se verificar o comportamento do processo produtivo. As figuras 35 e 36 mostram as condições de controle para as observações. A figura 35 mostra os gráficos da média (\bar{X}) e da amplitude (R) para os dados originais.

FIGURA 35 - Gráficos da \bar{X} e da R para os dados observados.

Através da figura 35 observa-se a sequência de observações e os limites dos gráficos tradicionais de Shewhart, onde encontra-se dois valores fora dos limites de controle para o gráfico da (\bar{X}), indicando que o processo não está sob de controle. Como podemos observar, os dados também aparentam exibir um comportamento sistemático.

A figura 36 mostra os gráficos da média (\bar{X}) e da amplitude (R) para os dados transformados. Após o ajuste dos dados, houve uma ação para corrigir o problema. O problema estava mascarado pela observações, a condição de normalidade foi verificada, e os gráficos de controle \bar{X} e R foram refeitos.

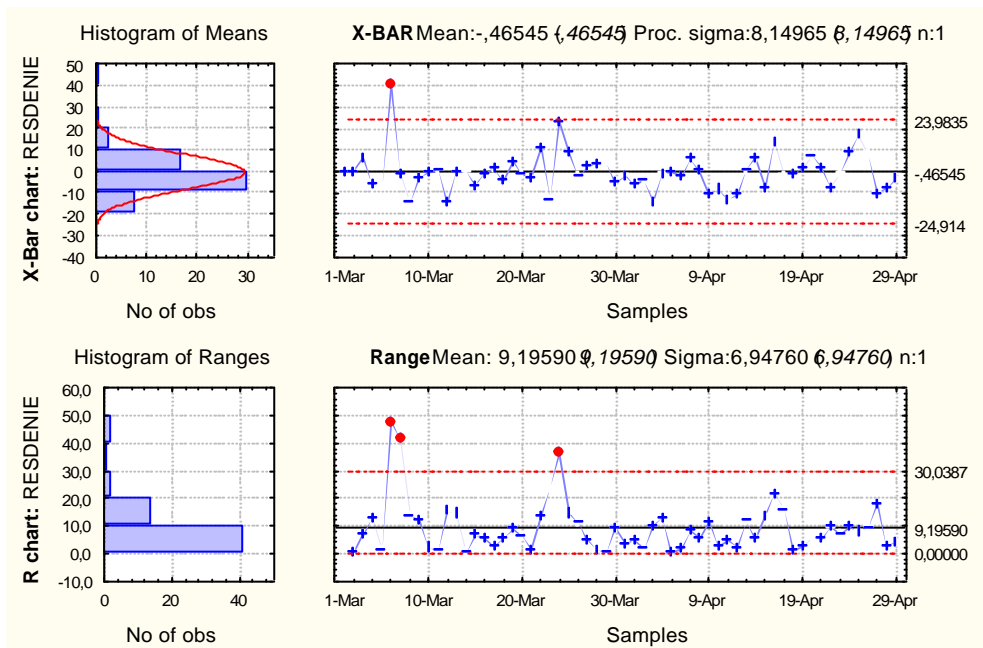


FIGURA 36 - Gráficos da \bar{X} e da R para os dados transformados.

Através da figura 36, podemos observar que os gráficos de controle para os mesmos dados, indicam que a observação 24 aparece, aqui, dentro dos limites de controle para a média. Isto é inteiramente possível em gráficos de controle tradicionais; porque os pontos estão fora dos limites pela causa da sistemática ou das causas comuns e não por causa de ocorrência de causas especiais.

3.3.2 Análise da série não-conforme no setor de tecelagem – ano 2002

O objetivo desta etapa é aplicar as técnicas descritas na seção 2.4, página 28. A análise segue os passos do fluxograma da figura 5, descrita no item 3.1, página 44. Os dados da tabela 14 são os valores diários da verificação da conformidade no setor de tecelagem da fita de polipropileno.

TABELA 14 - Valores diários coletados do número de itens não conformes no setor de tecelagem no período de 01 de março a 30 de abril de 2002

Data	Valores									
Março										
1 a 10/mar	6	1	8	5	3	8	1	5	4	2
11 a 20/mar	10	6	3	3	3	4	3	2	2	5
21 a 30/mar	4	3	5	7	2	3	7	5	2	1
31/mar	2									
Abril										
1 a 10/abr	3	4	5	8	7	3	2	4	1	4
11 a 20/abr	3	3	0	3	3	4	4	3	3	3
21 a 30/abr	2	3	1	5	3	2	3	3	3	2

Na figura 37, mostra-se a série do número diário de não conformidade no setor da tecelagem, no período de março a abril de 2002, num total de 61 valores. Observa-se que a série, aparentemente apresenta ser aleatória, não apresentando uma tendência.

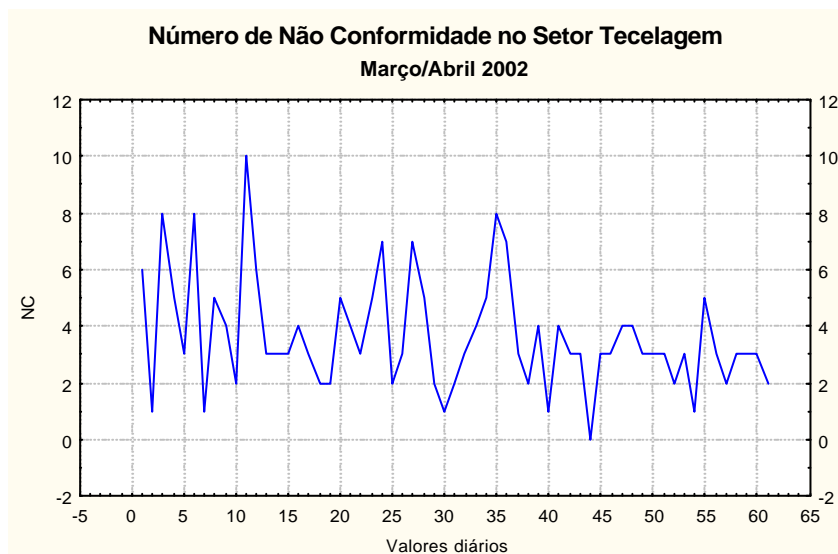


FIGURA 37 – Contagem diária dos dados

Pela figura 37, podemos observar que os dados possuem uma grande variabilidade. Os dados, aparentemente, não apresentam uma tendência sobre o tempo, o que iremos confirmar mais tarde. Para a verificação do tipo de distribuição, o teste de Kolmogorov-Smirnov foi executado:

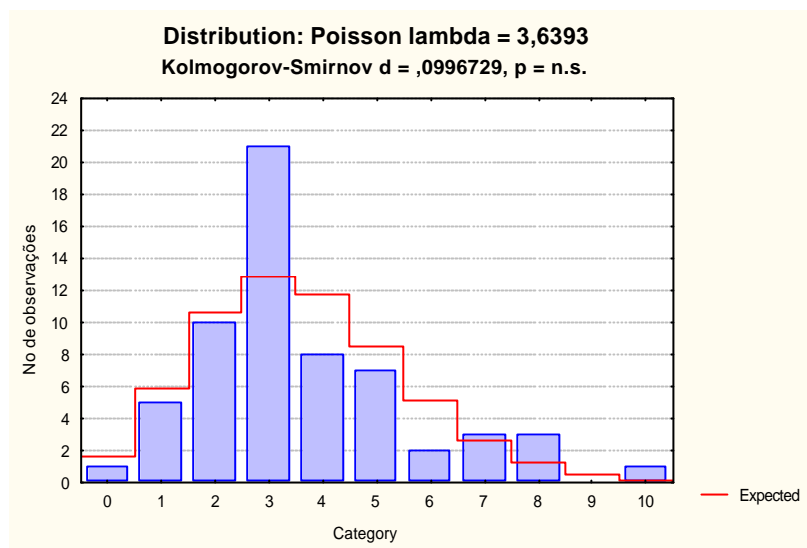


FIGURA 38 – Teste de aderência

Pela figura 38, verifica-se através do teste de Kolmogorov-Smirnov que as observações seguem uma distribuição de Poisson. A média de não conformidade é de 3,64 e o desvio padrão 2,00. Foi feito o ajustamento dos dados através do modelo de regressão de Poisson, onde encontramos os seguintes resultados:

TABELA 15 – Sumário dos parâmetros do modelo

	coef	Erro Padrão (SE)	coef/SE	valor p
Z	1.29	0,067	19,25	0

O algoritmo de Newton-Raphson convergiu em 4 iterações.

O modelo de regressão de Poisson encontrado foi:

$$Z = 1,29 + erro_{(0,067)}$$

Para verificar a adequação do modelo encontrado, usa-se os testes de Pearson e o desvio G^2 . A tabela 16 contém informações sobre a avaliação do modelo:

TABELA 16 - Critérios de avaliação do modelo

Critérios	GL - Graus de Liberdade	Valores	Valores/GL
Escala " Deviance" (G^2)	59	58,125	0,985
Escala de Pearson (χ^2)	59	55,885	0,947
Verossimilhança	59	64,78	

A Estatística de Desvio (G^2) e de Pearson (c^2) divididos pelos respectivos graus de liberdade são usados para detectar se há superdispersão ou subdispersão. A média e a variância numa distribuição de Poisson são iguais, isso implica que a Estatística de Pearson (c^2) dividida pelo grau de liberdade deve ser aproximadamente 1 (um). Pelos valores da tabela 16, verifica-se que o modelo é adequado aos dados.

Examinando os valores encontrados no gráfico da figura 39, observa-se que o comportamento dos dados indica adequação do modelo. Como as escalas das estimativas dos GLM não são sempre independentes, um pouco mais de cautela deve-se ter na interpretação do gráfico da Half-Normal (Lewis; Montgomery; Myers, 2001).

O valor residual do Desvio (G^2) indica a ausência de “outliers” (valores discrepantes). O gráfico da probabilidade meio-normal dos desvios residuais, figura 39, também confirma a ausência de “outliers”.

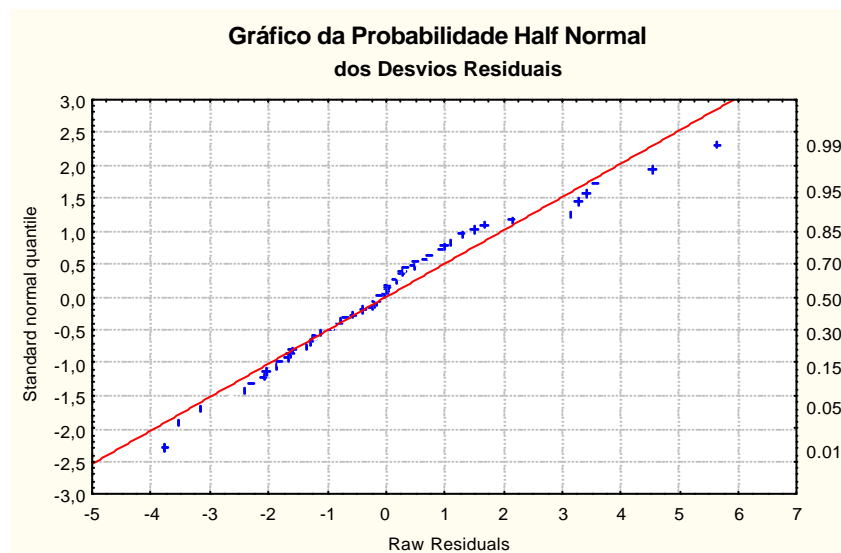


FIGURA 39 – Probabilidade half normal dos desvios residuais

O modelo encontrado é adequado. Segundo Piegorsch (1998), a análise gráfica e outras operações sobre os resíduos, ajudam a confirmar a adequação do modelo potencial corrente. Box plots dos resíduos do desvio (G^2), dos resíduos de

Pearson (χ^2) e o gráfico dos resíduos, podem confirmar o ajuste ou podem identificar efeitos associados com a variável de previsão.

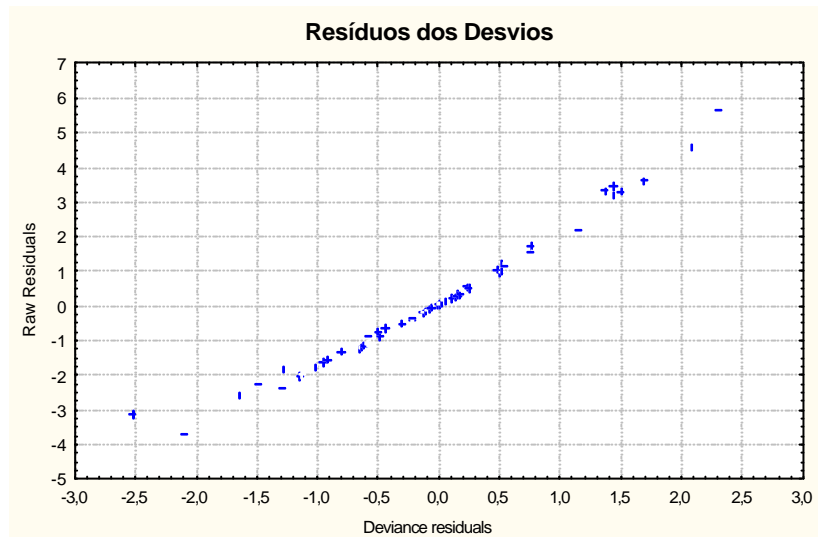


FIGURA 40 – Gráfico dos resíduos do desvio (G^2)

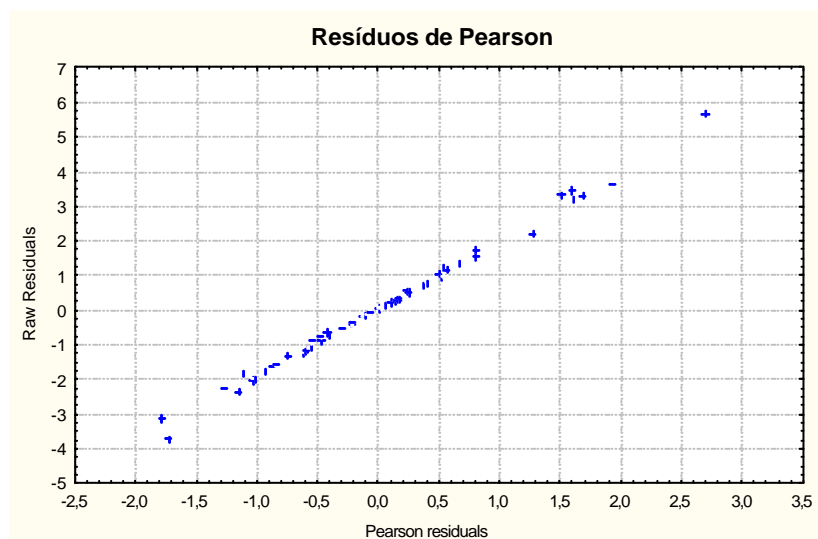


FIGURA 41 – Gráfico dos resíduos de Pearson

Os gráficos das figuras 40 e 41, mostram que ambos os tipos de resíduos são equivalentes para as observações, o que equivale dizer que os resíduos do modelo de regressão de Poisson se apresentam relativamente estáveis (Piegorisch, 1998).

Aplicando os gráficos de controle

Agora, pode-se verificar o comportamento do processo produtivo. As figuras 42 e 43 mostram as condições de controle para as observações. A figura 42 mostra o gráfico (U) para os dados originais.

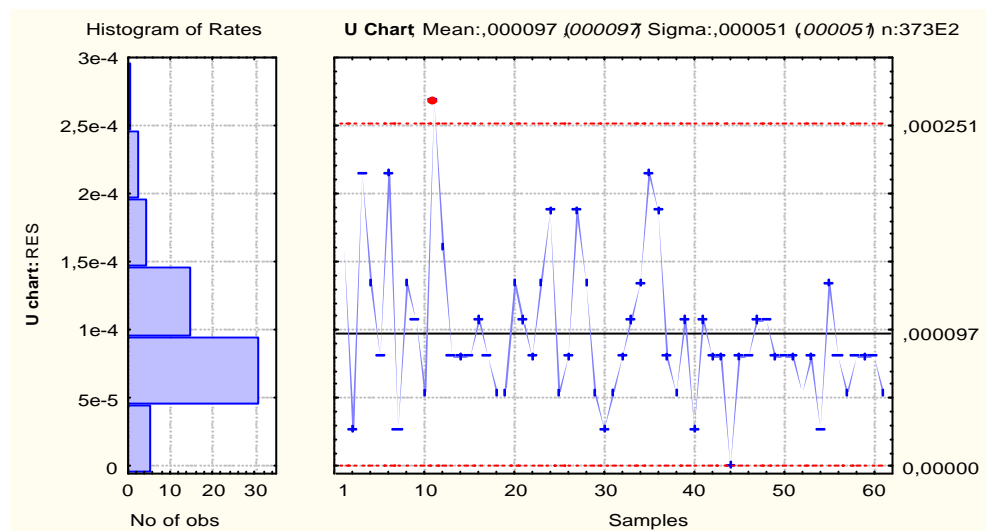


FIGURA 42 – Gráfico U para os dados observados.

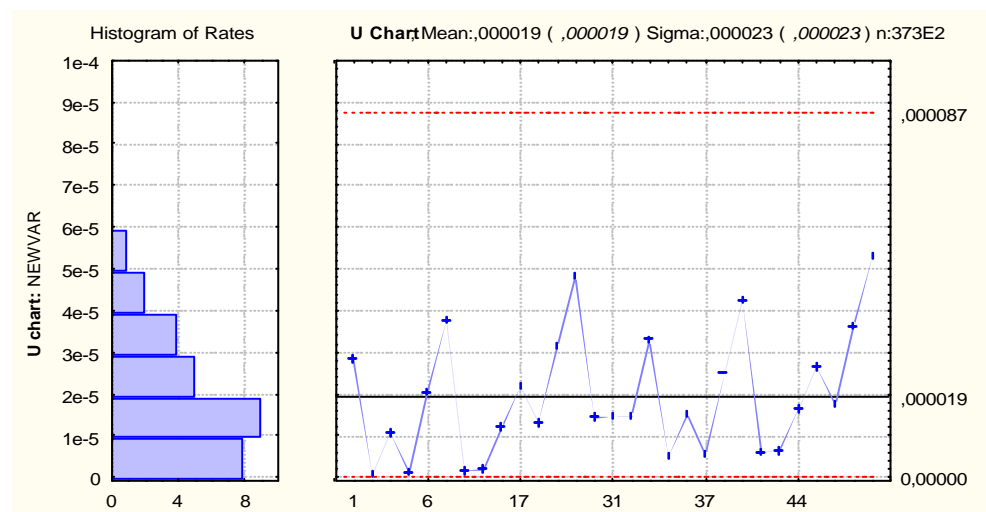


FIGURA 43 – Gráfico U para os dados modelados.

O sistema não estava ajustado e foi refeito o gráfico de controle U utilizando-se os resíduos encontrados pelo modelo de regressão de Poisson para corrigir o problema. Havia um ponto fora dos limites de controle por causa da sistemática ou das causas comuns e não por causa de ocorrência de causas especiais.

3.3.3 Ajuste final – ano de 2002

Nesta etapa, pretende-se ligar as duas etapas anteriores (subseção 3.3.1, página 65; e a subseção 3.3.2, página 72), de maneira a se conseguir detectar o número possível de não conformidade no resultado do processo. Para isso, aplica-se o teste de causalidade e ajusta-se um modelo via função de transferência, descrita na seção 2.5, página 34; utilizando-se como variáveis de entrada os resíduos obtidos através da regressão de Poisson e dos modelos Box e Jenkins.

Ao se analisar os resíduos obtidos através da regressão de Poisson, observa-se que existe uma correlação positiva no lag 4 e 12 do coeficiente de autocorrelação (figura 44), e no lag 4 do coeficiente de autocorrelação parcial (figura 45), indicando que se pode ajustar um modelo autorregressivo.

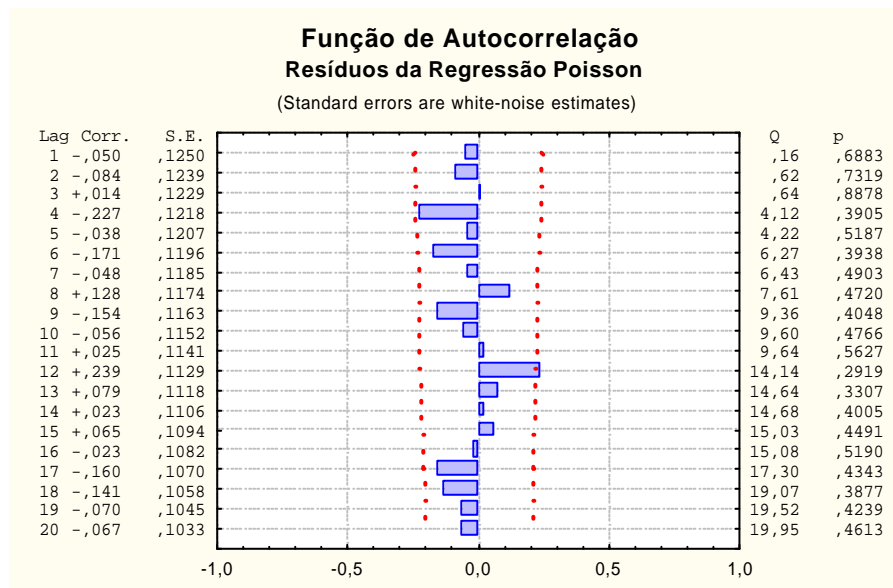


FIGURA 44 – Gráfico da função de autocorrelação

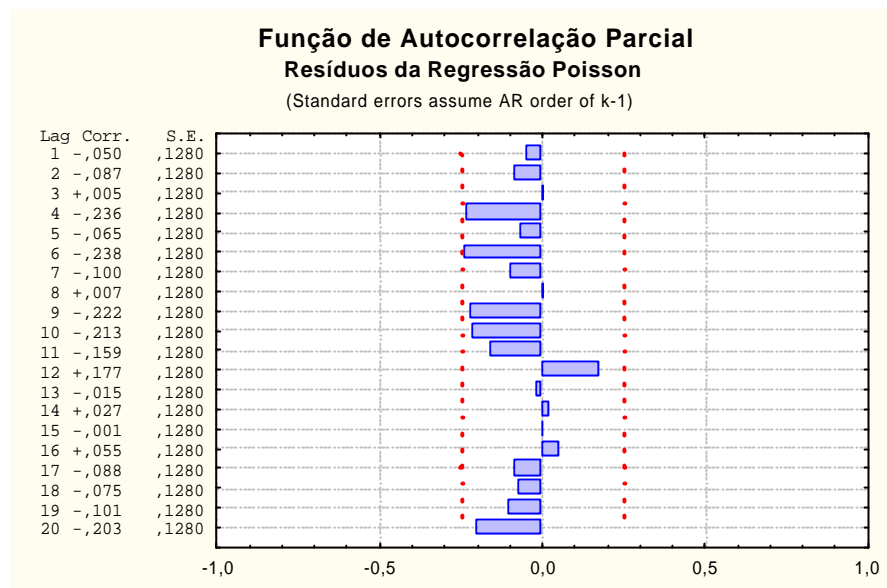


FIGURA 45 – Gráfico da função de autocorrelação parcial

Ao aplicar-se o teste de causalidade descrita na seção 2.5.2, página 38, pode-se afirmar que existe uma relação de causalidade unidirecional instantânea ou nível de significância de 5%, ou seja, a correlação cruzada é significativa para os lag 4, assim ($\rho(4) = 0,290$).

Através da tabela 17, observou-se que a variável gramatura do fio é uma variável explicativa da variável tecido.

TABELA 17 - Critérios de avaliação do modelo

Fator	Lag	Coef.	Tcal
Média		4,172	22,22
Autorregressivo	3	0,6707	7,66
Gramatura	6	0,0130	2,45

O modelo encontrado foi:

$$Z = 4,172 + 0,6707Z_{t-3} - 0,0130X_{t-6} + a_t$$

Aplicando os gráficos de controle

Agora se pode verificar o comportamento do processo produtivo. A figura 46, página 83, mostra as condições de controle para as observações modeladas através

da função de transferência e, mostra o gráfico (U) para os dados obtidos com os resíduos da função de transferência.

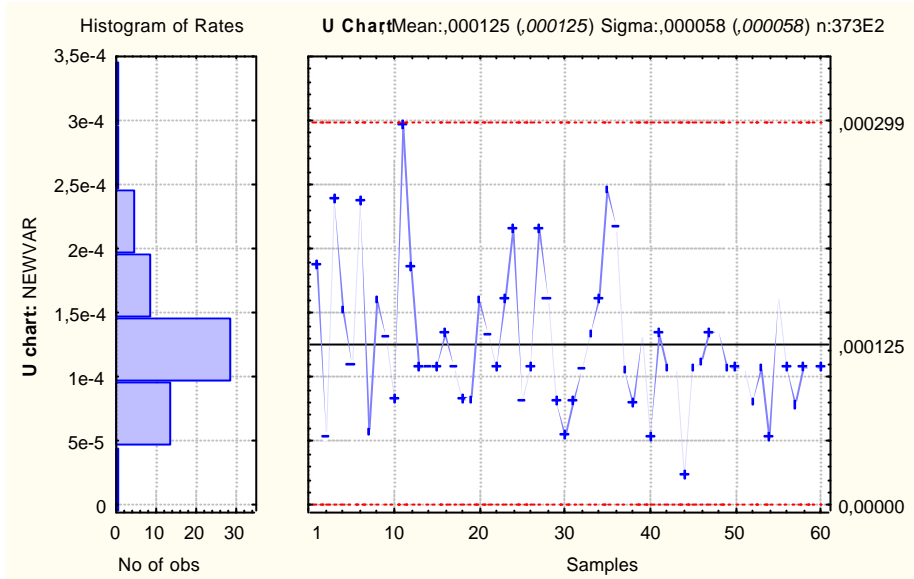


FIGURA 46 – Gráfico (U) para os dados obtidos com os resíduos da função de transferência.

Neste gráfico, figura 46, consegue-se detectar que nenhuma amostra está fora dos limites de controle, o que anteriormente não aparecia, pois estava mascarada pela autocorrelação, sendo assim o modelo encontrado através da função de transferência é mais sensível aos dados.

3.3.4 Comentários da análise dos dados – ano 2002

Verificou-se que, a técnica de se estudar a autocorrelação dos dados utilizada na análise dos dados da seção 3.2, mostrou-se satisfatória no estudo dos dados da seção 3.3. Os procedimentos propostos de analisar-se primeiro as variáveis contínuas (explicativas) e as variáveis discretas (independentes), e depois uni-las através da função de transferência foi crucial para confirmar a causalidade da gramatura da fita de rafia na tecelagem. Com isso pode-se contribuir para uma correta tomada de decisão na empresa.

3.4 COMENTÁRIOS GERAIS DO CAPÍTULO

Neste capítulo procurou-se apresentar uma aplicação da teoria utilizando-se de dados reais. Foram apresentados os procedimentos para se usar as técnicas de

controle estatístico de qualidade em observações correlacionadas. Além de se verificar a autocorrelação dos dados, pode-se modelar variáveis discretas através de modelos de regressão de Poisson e as variáveis contínuas através dos modelos Box e Jenkins, e com os resíduos obtidos utilizar os modelos de função de transferência para se identificar a existência de causalidade.

IV – CONCLUSÃO E RECOMENDAÇÕES

Este trabalho procurou apresentar uma metodologia alternativa às técnicas tradicionais de Controle Estatístico de Qualidade, através dos gráficos de controle modificados para análise de variáveis não-conformes autocorrelacionadas de um sistema produtivo. Para tal, foram coletados dados na Indústria Têxtil Oeste Ltda. no período de julho a agosto de 2001 e fevereiro a abril de 2002. Nos dados coletados foram analisados, simultaneamente, as variáveis contínuas e as variáveis discretas, para se encontrar uma das possíveis causas da falta de controle no processo final da produção. Foram apresentados os métodos para se usar as técnicas de controle estatístico de qualidade em observações correlacionadas. Verificou-se a autocorrelação dos dados, modelando-se as variáveis discretas através de modelos de regressão de Poisson e as variáveis contínuas através dos modelos Box e Jenkins. Com os resíduos obtidos nos modelos, utilizou-se os modelos de função de transferência para se identificar a existência de causalidade e fazer a combinação entre as variáveis.

Conclui-se que o procedimento proposto, de primeiro retirar a autocorrelação dos dados para depois ajustá-los, mostrou-se satisfatória estatisticamente. Ao estudar a autocorrelação dos dados temos uma nova perspectiva de aprendizagem sobre o processo produtivo através das informações contidas na estrutura de autocorrelação, dos modelos Box e Jenkins e dos modelos de regressão de Poisson, os quais eram ignorados pelo modelo clássico de monitoramento. A função de transferência empregada posteriormente nos resíduos obtidos permitiu a confirmação da causalidade da gramatura da fita de ráfia com relação à não conformidade do tecido produzido.

A titulação da gramatura da fita de ráfia é verificadas suas especificações no início do processo do sistema de produção da fábrica. A não conformidade do tecido produzido é verificada suas especificações no final do sistema de produção da fábrica. Antes deste estudo, supunha-se que a gramatura da fita de ráfia influenciava na fabricação do tecido, hoje se conseguiu provar que, realmente, a

gramatura da fita de ráfia é uma variável explicativa da não conformidade na fabricação do tecido.

Com este estudo, verificou-se que, quando a gramatura da fita de ráfia estiver fora das especificações, um a dois dias após, têm-se itens não conformes na fabricação do tecido. Assim, com o modelo estatístico desenvolvido, pode-se fornecer informações para que pessoas envolvidas diretamente no processo, possam realmente alcançar a melhoria contínua no processo.

Foram utilizados dados diários, e o procedimento proposto neste trabalho vem a ser a maior contribuição desta tese. Muitas empresas, por acreditarem nas vantagens que podem ser obtidas com a prática de CEQ, investem muitos recursos na implantação, principalmente, dos gráficos de controle convencionais, denominados gráficos de Shewhart. Como não é necessário um conhecimento profundo de estatística, é mais favorável a implantação desses gráficos pelas empresas, mas nem sempre os resultados são os esperados. Não há uma preocupação com a correlação dos dados. Neste contexto, o texto apresentado ao longo desta tese pode servir de referência às indústrias que encontram dificuldades em implantar o controle estatístico de qualidade. Entretanto, deve-se tomar cuidado com o tipo de variáveis a que está se propondo analisar, o que nos permite concluir que esta proposta de combinação das variáveis contínuas e discretas, pretende ser completa e extensiva a cobrir todas as possíveis dificuldades que possamos encontrar. No modelo de monitoramento clássico, não existe esta informação, ao identificar um item não conforme (variável discreta) no final do processo, não se sabe como proceder para que a mesma não ocorra, pois as variáveis utilizadas nos processos anteriores são contínuas.

Neste trabalho, sustenta-se que os modelos de séries temporais e os modelos de regressão de Poisson podem ser utilizados conjuntamente através de seus resíduos num modelo de função de transferência para ajustar as variáveis em estudo da melhor forma possível, visando com isso a um crescimento de informações para a correta tomada de decisão e pode-se detectar que houve uma melhora nos pontos de saída de controle.

Dessa maneira, a metodologia proposta permite um controle eficaz, possibilitando baixar os custos, pois o número de itens defeituosos produzidos poderá ser reduzido, permitindo que a Indústria seja mais competitiva.

5.1 SUGESTÕES PARA TRABALHOS FUTUROS

Como sugestões para trabalhos futuros ligados ao tema, pode-se destacar os relacionados a seguir, os quais abrangeriam áreas ainda com grandes lacunas a serem pesquisadas para um melhor entendimento e compreensão dos estudos sobre CEQ:

- Utilização da análise de intervenção, no caso dos dados possuírem *outliers*;
- Utilização de outras metodologias de séries temporais, tais como: à metodologia estrutural de Harvey (1990) ou a metodologia Bayesiana de Harrison e West (1989).
- Utilização de gráficos de controle não paramétricos.
- Utilização de dados com longa dependência.
- Utilização de outras funções de ligações.
- Avaliar a capacidade de processos. Enquanto alguns autores vêm com bons olhos a utilização de índices para a avaliação da capacidade de processos, outros têm questionado a validade desses em função das restrições práticas e de ordem estatística acerca do tema.

V – REFERÊNCIAS BIBLIOGRÁFICAS

ALWAN, L.C. **An investigation of the pervasiveness of control chart violations.** Management Research Center Working Paper. University of Wisconsin - Milwaukee pp. 90-01. 1990.

ALWAN, L. C.; ROBERTS, H.V. **Times series modeling for statistical process control.** Journal of Business Economics Statistics v16, pp. 83-89. 1988.

ALWAN, L. C. **Autocorrelation: fixed versus variable control limits.** Quality and Reliability Engineering International. v4 (2). pp. 167-188. 1991.

BARNARD, G. A. **Control charts and stochastic process.** Journal of the Royal Statistical Society, Series B, v.21, pp. 239-271. 1959.

BARTLETT, M. S. **On the theoretical specification of sampling proprieties of autocorrelated time series.** Journal of the Royal Statistical Society, Series B, v.8, pp. 27-41. 1946.

BAXLEY, R. V. Jr. **Discussion.** Technometrics, v.16. pp. 13-16. 1990.

BELTRÃO, K. I. **Séries temporais no domínio da frequência: uma introdução.** 4ª Escola de Séries Temporais e Econometria UFRJ/RJ. 1991.

BÖCKENHOLT, U. **Mixed INAR(1) Poisson regression models: analyzing heterogeneity and serial dependencies in longitudinal count data.** Journal of Econometrics. v.89 pp.317-338. 1999.

BOURKE, P.D. **Detecting a shift in fraction nonconforming using run-length control charts with 100% inspection** Journal of Quality Technology, v.23, pp. 225-238. 1991.

BOX, G. E. P; JENKINS, G. M. **Times series analysis: forecasting and control,** San Francisco: Holden-Day, 1970.

BOX, G. E. P; JENKINS, G. M. **Times series analysis: forecasting and control**, 2 ed. San Francisco: Holden-Day. 1976.

BOX, G. E. P; JENKINS, G.M.; REINSEL, G. C. **Times series analysis: forecasting and control**, 3 ed. San Francisco: Holden-Day. 1994.

BOX, G.E.P; LUCENO, A. **Statistical control by monitoring and feedback adjustment**, New York: Wiley & Sons. 1997.

BRAGA, A.L.F.; SALDIVA, P.H.N.; PEREIRA, L.A.A.; MENEZES, J.J.C.; CONCEIÇÃO, G.M.S.; LIN, C.A; ZANOBETTI, A.; SCHWARTZ, J.; DOCKERY, D.W. **Health effects of air pollution exposure on children and adolescents in São Paulo, Brazil**. Pediatric Pulmonology. V.31, 106-113. 2001.

BRESLOW, N. E. **Extra-Poisson variation in log-linear models**. Applied Statistical. V.33 pp. 38-44. 1984.

BROCKWELL, P. J. ; Davis, R. A. **Times series: theory and methods**. Springer-Verlag, Nova York. 1987.

CHRISTENSEN, R. **Log-linear models and logistic regression**. 2nd Edition. Springer-Verlag. Chapter 9. Generalized linear Models. 1997.

CONCEIÇÃO, M.S.C.; MIRAGLIA, S.G.E.K.; KISHI, H.S.; SALDIVA, P.H.N.; SINGER, J.M. **Air Pollution and child mortality: a time-series study in São Paulo, Brazil**. Environmental Health Perspectives, v.109 (suppl 3), 347-350. 2001.

CORDEIRO, G. **Introdução à teoria de verossimilhança**. Livro Texto do 10º Simpósio Nacional de Probabilidade e Estatística. UFRJ/ABE. Rio de Janeiro. 1992.

COX, D. R. **Statistical analysis of time series, some recent developments**. Scand. Journal Statistical v.8 pp. 93-115. 1981.

COX, D. R. **Some remarks on over-dispersion**. Biometrika. v.70 pp.267-274. 1983.

DEAN, C. **Testing for overdispersion in Poisson and statistical regression models**. Journal of the American Statistical Association v.83 pp. 173-178. 1992.

DEAN, C.; LAWLESS, J. F. **Tests for detecting overdispersion in Poisson regressions models**. Journal of the American Statistical Association. V84, n.406 pp. 467-471. 1989.

DEMING, W. E. **Out of the crises**. Massachusetts. Institute of Technology, Center of Advanced Engineering Studies. Cambridge. MA. 1986.

DILLON, W. R.; KUMAR, A. **Latent structure and other mixture models in marketing: an integrative survey and overview**. In: Bagozzi R. P. (Ed) Advanced Methods of Marketing Research Blackwell. Cambridge. Pp.352-388 1994.

DOBSON, A. J. **An introduction to generalized linear models**. 2 ed. CHAPMAN & HALL / CRC Press. Pp.225. 2002.

ENDER, P. **Applied categorical & nonnormal data analysis: Poisson models**.
 Class Notes UCLA California. 2002.
<http://www.gseis.ucla.edu/courses/ed231c/notes1/pois1.html>

FERRARI, S. L. P.; DAVID, J.S. E.; ANDRÉ, P. A.; PEREIRA, L.A. A. **Use of overdispersed regression models in analyzing the association between air pollution and human health**. Relatório Técnico, RTMAE-2002-10, IME-USP. 2002

GARDINER, J. S.; MONTGOMERY, D. C. **Using statistical control charts for software quality control**. Quality and Reliability Engineering International, v.3. 1987.

GRANGER, C. W. **Investigating causal relationships by econometric models and cross-spectral methods**. Econometrica v. 37 pp. 424-438. 1969.

GRANGER, C. W. e MORRIS, M. J. **Time series modeling and interpretation**. Journal of the Royal Statistical Society, Series A, v.139 pp. 246-257. 1976.

GRIFO, E. **Controle estatístico do processo**. ABIMAQ. 1997.

HARRIS, T. J.; ROSS, W. H. **Statistical process control procedures for correlated observations**. Canadian Journal of Chemical Engineering, v.69, pp.48-57. 1991.

HARRISON, P. J.; WEST, M. **Bayesian forecasting and dynamic models**. London: Springer-Verlag London. 1989.

HARVEY, A. C.; TODD, P. H. J. **Forecasting, structural time series models and Kalman filter**. Cambridge University Press, 2ed. 1990.

HINDE, J.; DEMÉTRIO, C. G. B. **Overdispersion: models and estimation**. 13^o Simpósio Nacional de Probabilidade e Estatística. Caxambu: UFMG/ABE. 1992

HUNTER, J. S. **Discussion**. Technometrics v.32. pp. 21-22. 1990.

JACKSON, J. E. **All count distributions are not alike** Journal of Quality Technology v.4. 1972.

JOHNSON, R.A.; BAGSHAW, M. **The effect of serial correlation on the performance of Cusum test**, Technometrics, v. 16, pp. 103-112. 1974.

JOHNSON, N.L.; KOTZ, S. **Discrete distributions**. Houghton Mifflin, Boston, Mass. 1969.

JING, S.; GONGXU, Z. **Control charts based on the number of consecutive conforming items between two successive nonconforming items for the near zero-nonconformity processes** Total Quality Management Abingdon v.11 pp. 235-250. 1993.

JURAN, J. M. **Juran na liderança pela qualidade**. 2^a Ed. São Paulo: Livraria Editora Pioneira. 1993.

LAGO NETO, J. C. **O efeito da autocorrelação em gráficos de controle para variável contínua: um estudo de caso**. Dissertação de mestrado em Engenharia de Produção e Sistema. UFSC. Florianópolis. SC. 1999.

LEWIS, S.; MONTGOMERY, D.C.; MYERS, R.H. **Examples of designed experiments with nonnormal responses**. Journal of Quality Technology V33 n.3 pp.265-278. 2001.

LUCAS, J. M.; SACCUCCI, M. S. **Exponentially weighted moving average control schemes: properties and enhancements**. Technometrics. v.32. (1) pp.1-12. 1990.

LOURENÇO FILHO, R. de C. B. **Controle estatístico de qualidade**. Ed. Livros Técnicos e Científicos. 1985.

MACGREGOR, J. F.; HARRIS, T. J. **Discussion**. Technometrics v.32 pp.23-26. 1990.

MARQUARDT, D. W. **New technical and educational directions for managing product quality**. The American Statistician USA. pp.8-14. 1984.

McCULLAGH, P. ; NELDER, J. A. **Generalized Linear Models**. Third Edition. New York: Chapman and Hall/CRC. Reprint. 2000.

MESSINA, W. **Statistical quality control for manufacturing manages**. New York: Wiley Series in Engineering Management. 1987.

MILLS, T. M.; SENETA, E. **Independence of partial autocorrelations for a classical immigration branching process**. Stochastic Processes and Their Applications v. 37 pp.275-279. 1991

MONTGOMERY, D. C. **Introduction to statistical quality** 3ed. v.2, USA: John Wiley e Sons. 1997.

MONTGOMERY, D. C.; FRIEDMAN, D. J. **Statistical process control in computer integrated manufacturing environment: in statistical process control for automated manufacturing** eds. J.B. Keats and N.F. Hubele, New York, N.Y.: Marcel Dekker, 1989.

MURTEIRA, B. J. F.; MULLER, D. A.; TURMKMAN, K. F. **Análise de sucessões cronológicas**. Portugal: Ed. McGraw-Hill. 1993.

NELDER, J. A.; WODDERBURN, R. W. M. **Generalized linear models**. Journal of Royal Statistical Society: v135 pp 370-384. 1972.

PAULA, G. A. **Estimação e testes em modelos de regressão com parâmetros restritos**. Livro Texto da 5ª Escola de Modelos de Regressão. IME-USP/ABE. Campos do Jordão. 1997

PAGE, E. S. **Continuous inspection schemes**. Biometrika, v.41 pp.100-114. 1954.

PIERGORSCH, W. W. **An introduction to binary response regression and associated trend analyses.** Journal of Quality Technology v.30 n.3 pp.269-281. 1998.

PINDYCK, R. S.; RUBINFELD, D. **Econometric models and economic forecasts.** Singapore. Ed. McGraw-Hill. 1991.

RAMOS, A. W. **Controle estatístico de processo para pequenos lotes.** São Paulo: Editora Edgard Blücher LTDA. 1995.

RAMOS, A. W. **Controle estatístico de processo para pequenos lotes.** São Paulo: Editora Edgard Blücher LTDA. 2000.

ROBERTS, S. W. **Control charts tests based on geometric moving averages** Technometrics, v.1, pp. 239-250. 1959.

RYAN, T. P. **Statistical methods for quality improvement.** Texas A&M University. USA: John Wiley & Sons, pp.71-73. 1989.

RUSSO, S. L. **A demanda de gasolina e óleo diesel no Brasil no período de 1980-1987.** Monografia de Especialização. Santa Maria: UFSM. 1989.

SAMOHYL, R. W.; RAMOS, E. M. L.; ALMEIDA, S. S. **Modelando a Heterogeneidade da Variância através do Glim.** IV Simpósio de Pesquisa Operacional da Marinha. V Simpósio de Logística da Marinha, 2001, Rio de Janeiro: SPOLM, 2001, v. 1, n. , pp. 1-1

SAMOHYL, R. W.; RUSSO, S. L. **Gráficos de Controle ARMA para Dados Correlacionados.** 3º Congresso Nacional de Pesquisadores. São Paulo: 3º CONAPE/ASSER. v. 1. pp. 115-115. 2001.

SAMOHYL, R. W.; RUSSO, S. L. **Modelos de regressão de Poisson para dados não-conformes aplicados na indústria.** 15º Simpósio Nacional de Probabilidade e Estatística. Águas de Lindóia: USP/ABE. 2002

SANTOS, A. R. **Metodologia científica: a construção do conhecimento.** 2 ed Rio de Janeiro: DP&A Editora. 1999

SCHAFER, J. L. **Analysis of incomplete multivariate data**. London Chapman & Hall. 1997.

SCHEAFFER, R. L.; LEAVENWORTH, R. S. **The negative binomial model for counts in units of varying size**. Journal of Quality Technology v.8. 1976.

SHEWHART, W. A. **Economic control of quality of the manufactured product**. Van Nostrand, New York. 1931.

STOKES, M.; KOCH, G. **A macro for maximum likelihood fitting of log-linear models to Poisson and multinomial counts**. Proceedings of the Eighth Annual SAS Users Group International; Cary. North Carolina: SAS Institute. 1983.

SOUZA, R. C.; CAMARGO, M. E. **Análise e previsão de séries temporais: os modelos ARIMA**. Ijuí: SEDIGRAF. 1996.

TRIOLA, M. F. **Introdução à estatística**. 7ª ed., Rio de Janeiro: Editora LTC . 1999.

VIEIRA, S. **Estatística para a qualidade: como avaliar com precisão a qualidade em produtos e serviços**. Rio de Janeiro: Editora Campus. 1999.

WANG, P.; PUTERMAN, M. L.; COCKBURN, I.; LE, N. **Mixed Poisson regression models with covariate dependent rates**. Biometrics. V. 52 pp. 381-400. 1996

WARDELL, D. G.; MOSKOWITZ, H.; PLANTE, R. D. **Run-length distributions of special-cause control charts for correlated processes**. Technometrics. v.36 n.1 pp. 3-17. 1994.

WEI, W. W. S. **Time series analysis – univariate and multivariate methods**. Addison-Wesley Publishing Company, Inc. 1990

WONG, C. M.; MA, S.; HEDLEY, A.J.; LAM, T. H. **Effect of air pollution on daily mortality in Hong Kong**. Environmental Health Perspectives, v.109, 335-340. 2001.

ZAHNER, G.E.P.; DASKALAKIS, C. **Factors Associated With Mental Health, General Health and School-Based Service Use for Psychopathology**, American Journal of Public Health, v.87, 1441-1448. 1997,

ZEGER, S. L. **A regression model for time series counts.** Biometrika. v.75, pp.621-629. 1988.

ZEGER, S. L.; LIANG, K-Y. **Longitudinal data analysis for discrete and continuous outcomes.** Biometrics. v. 42 pp.121-130. 1986

ZEGER, S. L.; QAQISH, B. **Markov regression models for time series; a quasi-likelihood approach.** Biometrics. v. 44 pp. 1019-1031. 1988.

ZHANG, N. F. **A statistical control chart for stationary process data.** Technometrics. V.40 pp.24-38. 1998.

ANEXO A - Modelos auto-regressivos

Será caracterizada a generalização para o AR(p), MA(q) e os Modelos Mistos (Murteira; Müller e Turmkman, 1993).

MODELO AR(p)

Modelo: $\tilde{Z}_t = f_1 \cdot \tilde{Z}_{t-1} + f_2 \cdot \tilde{Z}_{t-2} + \dots + f_p \cdot \tilde{Z}_{t-p} + a_t$ (A1)

Estacionaridade: as raízes de $f(B) = 0$ devem cair fora do círculo unitário.

Inversibilidade: como $p(B) \equiv f(B)$ é tal que $\{p_n\} < \infty$, o AR(p) é sempre inversível.

Média = 0

Variância: $s_z^2 = \frac{s_a^2}{(1 - f_1 r_1 - \dots - f_p r_p)}$ (A2)

Autocovariâncias: $f(B)g_k = 0 \quad \forall k > 0$, B operando em k (A3)

Autocorrelações: $f(B)r_k = 0 \quad \forall k > 0$, (A4)

então: $r_k = f_1 \cdot r_{k-1} + f_2 \cdot r_{k-2} + \dots + f_p r_{k-p}$ (A5)

Autocorrelações parciais: $f_{kk} = \frac{|R_k^*|}{|R_k|} \quad f_{kk} = 0 \quad \forall k > p$ (A6)

Padrões para identificação: têm-se 2 casos:

- 1) se as raízes de $f(B) = 0$ são reais, r_k cai exponencialmente a zero com o crescimento de k alternando ou não o sinal;
- 2) se as raízes são complexas, r_k é uma senóide amortecida.

Em geral, a função de autocorrelação de um AR(p) é uma mistura de exponenciais e ondas-seno amortecidas.

As autocorrelações parciais para um AR(p) são nulas para $k > p$.

MODELO MA(q)

Modelo: $\tilde{Z}_t = m + q(B)a_t$ (A7)

Estacionaridade: como $\{y_n\} < \infty \rightarrow$ o modelo MA(q) é sempre estacionário.

Inversibilidade: o modelo será inversível se as raízes de $q(B) = 0$ caírem fora do círculo unitário.

Variância: $s_z^2 = (1 + q_1^2 + q_2^2 + \dots + q_q^2)s_a^2$ (A8)

Autocovariâncias: $g_k = g_{az}(k) + q_1 g_{az}(k-1) + \dots + q_q g_{az}(k-p)$ (A9)

$$\text{onde } g_{az}(j) = \begin{cases} s_a^2 & j = 0 \\ 0 & j > 0 \\ q(B)g_{aa}(-j) & j < 0 \end{cases}$$

onde $g_{aa}(j) = 0 \quad \forall j \neq 0$

$$g_k = 0 \quad \forall k > q$$

Autocorrelações: $r_k = \frac{g_k}{g_0}$, (A10)

como $g_k = 0 \quad \forall k > q \rightarrow r_k$ sofre um corte após $k = q$.

Autocorrelações parciais: comportam-se como $\{r_k\}$ de um AR(p).

Padrões para identificação: a função de autocorrelação é nula para $k > q$.

1) $\{f_{kk}\}$ cai exponencialmente à medida que k aumenta, alternando ou não de sinal.

2) $\{f_{kk}\}$ é uma senóide amortecida.

MODELOS MISTOS

Para os modelos mistos, isto é, aqueles que contém ambos os operadores, tanto o auto-regressivo como o de média móvel, identificados como ARMA(p,q) será

particularizado somente o comportamento do ARMA (1,1). (Murteira; Müller e Turmkman, 1993)

MODELO ARMA (1,1)

Modelo: $\tilde{Z}_t = f z_{t-1} + a_t - q a_{t-1}$ (A11)

Estacionaridade: o modelo é estacionário se as raízes de $(B) = 0$ caem fora do círculo unitário, isto é $-1 < f < 1$

Inversibilidade: o modelo será inversível se as raízes de $q(B) = 0$ caírem fora do círculo unitário, isto é , $-1 < q < 1$

Pode-se então, construir a região de viabilidade do modelo em função dos parâmetros.

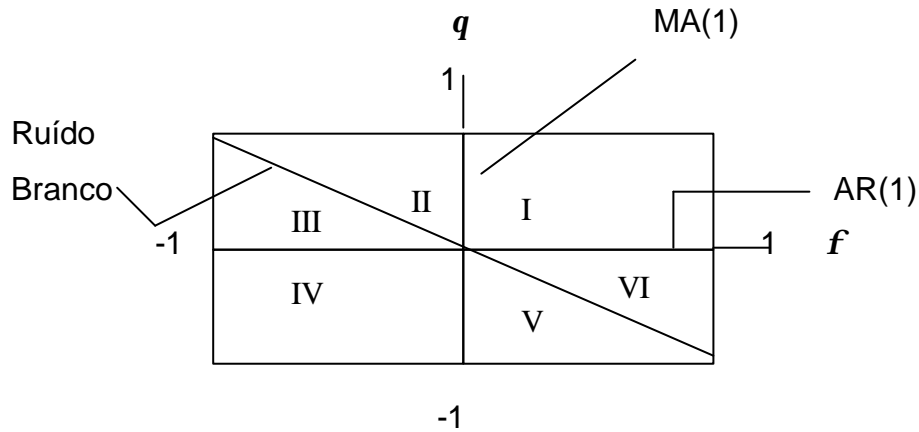


FIGURA A1 – Região admissível para o ARMA(1,1) em função dos parâmetros.

Variância: $g_x^2 = g_0$ (A12)

$$s_x^2 = \frac{1 + 2fq + q^2}{1 - f^2} s_a^2 \quad (A13)$$

Autocovariâncias: $g_0 = \frac{1 + q^2 - 2fq}{1 - f^2} s_a^2$ (A14)

$$g_1 = \frac{(f - q)(1 - fq)}{1 - f^2} s_a^2 \quad (A15)$$

$$\mathbf{g}_k = f\mathbf{g}_{k-1} \quad \forall k \geq 2 \quad (\text{A16})$$

$$\text{Autocorrelações: } \mathbf{r}_k = \frac{\mathbf{g}_k}{\mathbf{g}_0}. \text{ Assim,} \quad (\text{A17})$$

$$\mathbf{r}_1 = \frac{(f - q)(1 - fq)}{1 - 2fq + q^2} \quad (\text{A18})$$

$$\mathbf{r}_2 = f\mathbf{r}_1 \quad (\text{A19})$$

$$\mathbf{r}_k = f\mathbf{r}_{k-1} \quad \forall k \geq 2 \quad (\text{A20})$$

Pode-se assim, delimitar a região admissível para um ARMA(1,1) em função das autocorrelações de baixa ordem.

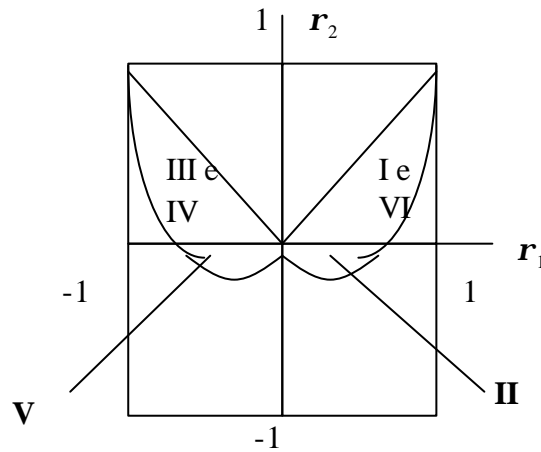


FIGURA A2 – Região admissível para o ARMA(1,1) em função das autocorrelações \mathbf{r}_1 e \mathbf{r}_2 .

Autocorrelações parciais: f_{kk} cai exponencialmente a zero, alternando ou não de sinal a partir de $f_{11} = ?_1$.

ANEXO B - Ajuste do modelo de regressão de Poisson através do algoritmo de Newton-Raphson

Com a ligação log, o procedimento do IRWLS (Iteratively Reweighted Least Squares) corresponde ao algoritmo Newton-Raphson que é dado por: (maiores detalhes ver Stockes e Koch, 1983)

$$\hat{\mathbf{b}} = (\mathbf{X}^T \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W} \mathbf{Z}, \quad (\text{B1})$$

$$\text{onde } \mathbf{W} = \text{Diag} \left[\text{Var}(y_i) \left(\frac{\partial \mathbf{h}_i}{\partial \mathbf{m}_i} \right)^2 \right]^{-1} \quad (\text{B2})$$

é a matriz dos pesos e $\mathbf{z} = \mathbf{h} + \left(\frac{\partial \mathbf{h}}{\partial \mathbf{m}} \right) (y - \mathbf{m})$ é a variável dependente ajustada de trabalho. Porque $\mathbf{h}_i = \log \mathbf{m}_i$, segue que $\partial \mathbf{h}_i / \partial \mathbf{m}_i = 1/\mathbf{m}_i$, assim,

$$\mathbf{W} = \text{Diag}(\mathbf{m}_i) \quad (\text{B3})$$

$$\mathbf{z} = \mathbf{h}_i + (y_i - \mathbf{m}_i) / \mathbf{m}_i \quad (\text{B4})$$

Dado a estimativa corrente para \mathbf{b} ,

- calcula-se $\mathbf{h}_i = \mathbf{x}_i^T \mathbf{b}$, $\mathbf{m}_i = \exp(\mathbf{h}_i)$ e
- $\mathbf{z}_i = \mathbf{h}_i + (y_i - \mathbf{m}_i) / \mathbf{m}_i$; então a regressão de \mathbf{z} sobre \mathbf{X} usando \mathbf{m}_i 's como pesos, obtêm-se a nova estimativa para \mathbf{b} ; e repete-se o cálculo até o valor de \mathbf{b} convergir (Schafer, 1999).

ANEXO C - Função desvio

È apresentado aqui, a demonstração da função desvio, que serve para testar a adequação de um modelo linear generalizado (GLM), que segundo Paula (1997), sem perda de generalidade, suponha que o log da função de verossimilhança seja definido por:

$$L(\mathbf{I} ; Z) = \sum_{i=1}^n L(I_i ; Z_i) \quad (C1)$$

onde $I_i = g^{-1}(\mathbf{h}_i)$ e $\mathbf{h}_i = x_i' \mathbf{b}$. Para o modelo saturado ($p = h$) a função $L(\mathbf{I} ; Z)$ é estimada por

$$L(Z, Z) = \sum_{i=1}^n L(Z_i ; Z_i) . \quad (C2)$$

onde Z_i é os valores observados das variáveis aleatórias independentes Z .

Note que, a estimativa de máxima verossimilhança de I_i fica, nesse caso, dada por $\hat{I}_i^s = Z_i$. Quando $p < h$, a estimativa de $L(\mathbf{I} ; Z)$ é denotada por $L(\hat{\mathbf{I}} ; Z)$. Aqui, a estimativa de máxima verossimilhança de I_i será dada por $\hat{I}_i = g^{-1}(\hat{\mathbf{h}}_i)$, onde $\hat{\mathbf{h}}_i = x_i^T \hat{\mathbf{b}}$.

A qualidade do ajuste de um GLM é avaliada através da função desvio

$$D^*(Z; \hat{\mathbf{I}}) = \text{fD}(Z; \hat{\mathbf{I}}) = 2\{L(Z; Z) - L(\hat{\mathbf{I}}; Z)\} , \quad (C3)$$

que é uma distância entre o log da função de verossimilhança do modelo saturado (com h parâmetros) e do modelo sob investigação (com p parâmetros) avaliado na estimativa de máxima verossimilhança irrestrita $\hat{\mathbf{b}}$. Um valor pequeno para a função desvio indica que, para um número menor de parâmetros, obtém-se um ajuste tão bom quanto o ajuste com o modelo saturado. Denotando por $\hat{\mathbf{q}}_i = \mathbf{q}_i(\hat{\mathbf{I}}_i)$ e $\hat{\mathbf{q}}_i^s = \mathbf{q}_i(\hat{\mathbf{I}}_i^s)$ as estimativas de máxima verossimilhança de \mathbf{q} para os modelos com

p parâmetros ($p < h$) e saturado ($p = h$), respectivamente, tem-se que a função $D(Z; \hat{I})$ fica, alternativamente, dada por:

$$D(Z; \hat{I}) = 2 \sum_{i=1}^n \{Z_i(\hat{q}_i^s - \hat{q}_i) + (b(\hat{q}_i) - b(\hat{q}_i^s))\}. \quad (C4)$$

Embora seja usual comparar os valores observados da função desvio com os percentis da distribuição qui-quadrado com $n - p$ graus de liberdade, em geral $D(Z; \hat{I})$ não segue assintoticamente uma χ_{n-p}^2 . No caso binomial quando k é fixo e $h_i \rightarrow \infty$ e $n_i I_i (1 - I_i)$ permanece limitado. Para o modelo de Poisson, quando $I_i \rightarrow \infty$ para todo i , tem-se $D(Z; \hat{I}) \sim \chi_{n-p}^2$. No caso normal, como é conhecido para o s^2 fixo, $D(Z; \hat{I}) \sim \sigma^2 \cdot \chi_{n-p}^2$. Como $E(\chi_r^2) = r$, o valor do desvio próximo de χ_{n-p}^2 pode ser uma indicação de que o modelo esta bem ajustado. Em geral, para os casos em, que $D^*(Z; \hat{I})$ depende do parâmetro de dispersão f^{-1} , o seguinte resultado (Jorgensen, 1987) para a distribuição nula da função desvio pode ser utilizado:

$$D^*(Z; \hat{I}) \sim \chi_{n-p}^2, \text{ quando } f \rightarrow \infty.$$

Isto é, quando a dispersão é pequena, fica razoável comparar os valores observados de $D^*(Z; \hat{I})$ com os percentis da χ_{n-p}^2 . Em particular, para o caso normal linear, o resultado acima diz que $\sum_{i=1}^n (Z_i - \hat{I}_i)^2 / s^2 \sim \chi_{n-p}^2$ quando $s^2 \rightarrow 0$. No caso do modelo gama, o desvio estará bem aproximado por uma distribuição de qui-quadrado com $n - p$ graus de liberdade a medida que o coeficiente de variação ficar próximo de zero.

ANEXO D - Normas

ANEXO E - Relação dos dados usados no trabalho

TABELA E1 - Cálculo da autocorrelação dos valores da gramatura da fita de polipropileno – ano 2001

	Valores	1	2	3	4	1x3	1x4
31-jul	685,46	20,34	413,7156	35,777	3,316	727,7123	67,4454
1-ago	700,90	35,78	1280,0224	3,316	14,470	118,6343	517,6882
2-ago	668,44	3,32	10,9952	14,470	8,931	47,9801	29,6153
3-ago	679,59	14,47	209,3722	8,931	13,393	129,2332	193,7898
4-ago	674,05	8,93	79,7681	13,393	5,906	119,6151	52,7447
5-ago	678,51	13,39	179,3671	5,906	9,008	79,0925	120,6450
6-ago	671,03	5,91	34,8761	9,008	15,752	53,1988	93,0238
7-ago	674,13	9,01	81,1477	15,752	14,060	141,8954	126,6508
8-ago	680,87	15,75	248,1192	14,060	7,418	221,4624	116,8547
9-ago	679,18	14,06	197,6695	7,418	9,418	104,3004	132,4194
10-ago	672,54	7,42	55,0341	9,418	4,162	69,8711	30,8765
11-ago	674,54	9,42	88,7081	4,162	7,572	39,2007	71,3197
12-ago	669,28	4,16	17,3231	7,572	11,649	31,5167	48,4851
13-ago	672,69	7,57	57,3397	11,649	-7,812	88,2112	-59,1571
14-ago	676,77	11,65	135,7039	-7,812	9,752	-91,0070	113,6007
15-ago	657,31	-7,81	61,0320	9,752	20,701	-76,1840	-161,7185
16-ago	674,87	9,75	95,0976	20,701	17,213	201,8671	167,8607
17-ago	685,82	20,70	428,5107	17,213	18,777	356,3239	388,7016
18-ago	682,33	17,21	296,2977	18,777	14,957	323,2210	257,4576
19-ago	683,90	18,78	352,5908	14,957	7,880	280,8517	147,9659
20-ago	680,08	14,96	223,7089	7,880	-3,082	117,8604	-46,0897
21-ago	673,00	7,88	62,0944	-3,082	3,290	-24,2822	25,9276
22-ago	662,04	-3,08	9,4956	3,290	-19,389	-10,1391	59,7478
23-ago	668,41	3,29	10,8261	-19,389	-19,864	-63,7963	-65,3572
24-ago	645,73	-19,39	375,9411	-19,864	-4,582	385,1393	88,8316
25-ago	645,26	-19,86	394,5626	-4,582	-14,248	91,0051	283,0205
26-ago	660,54	-4,58	20,9901	-14,248	-35,889	65,2781	164,4264
27-ago	650,87	-14,25	203,0112	-35,889	-1,197	511,3565	17,0537
28-ago	629,23	-35,89	1288,0347	-1,197	-1,710	42,9558	61,3598
29-ago	663,92	-1,20	1,4326	-1,710	-11,274	2,0463	13,4936
30-ago	663,41	-1,71	2,9231	-11,274	-0,530	19,2748	0,9067
31-ago	653,85	-11,27	127,0986	-0,530	-17,274	5,9785	194,7414
1-set	664,59	-0,53	0,2812	-17,274	-13,094	9,1603	6,9440
2-set	647,85	-17,27	298,3842	-13,094	-9,223	226,1900	159,3093
3-set	652,03	-13,09	171,4633	-9,223	-13,633	120,7644	178,5133
4-set	655,90	-9,22	85,0564	-13,633	22,495	125,7299	-207,4661
5-set	651,49	-13,63	185,8532	22,495	9,495	-306,6753	-129,4489
6-set	687,62	22,50	506,0430	9,495	-9,017	213,6028	-202,8500
7-set	674,62	9,50	90,1626	-9,017	-1,556	-85,6238	-14,7739
8-set	656,10	-9,02	81,3135	-1,556	-6,941	14,0302	62,5853
9-set	663,56	-1,56	2,4208	-6,941	-3,402	10,7987	5,2933
10-set	658,18	-6,94	48,1705	-3,402	-16,530	23,6123	114,7285

	Valores	1	2	3	4	1x3	1x4
11-set	661,72	-3,40	11,5743	-16,530	-2,787	56,2377	9,4806
12-set	648,59	-16,53	273,2508	-2,787	-9,735	46,0650	160,9291
13-set	662,33	-2,79	7,7657	-9,735	-5,248	27,1296	14,6252
14-set	655,38	-9,74	94,7780	-5,248	-2,556	51,0933	24,8827
15-set	659,87	-5,25	27,5436	-2,556	-12,633	13,4139	66,2995
16-set	662,56	-2,56	6,5326	-12,633	-16,812	32,2882	42,9706
17-set	652,49	-12,63	159,5876	-16,812	-0,556	212,3864	7,0226
18-set	648,31	-16,81	282,6534	-0,556	1,688	9,3460	-28,3741
19-set	664,56	-0,56	0,3090	1,688	-10,505	-0,9382	5,8395
20-set	666,81	1,69	2,8483	-10,505	0,059	-17,7286	0,1004
21-set	654,62	-10,50	110,3466	0,059	1,572	-0,6250	-16,5164
22-set	665,18	0,06	0,0035	1,572	-8,402	0,0936	-0,4999
23-set	666,69	1,57	2,4721	-8,402	1,188	-13,2106	1,8674
24-set	656,72	-8,40	70,5953	1,188	-6,376	-9,9792	53,5752
25-set	666,31	1,19	1,4106	-6,376	3,162	-7,5733	3,7556
26-set	658,74	-6,38	40,6585	3,162	-10,505	-20,1628	66,9815
27-set	668,28	3,16	9,9989	-10,505	-16,966	-33,2166	-53,6488
28-set	654,62	-10,50	110,3466	-16,966		178,2231	
29-set	648,15	-16,97	287,8519				
Soma	40572,40		10012,4861			4805,8832	3586,4570

TABELA E2 - Cálculo da autocorrelação dos valores da gramatura da fita de polipropileno – ano 2002

	Valores	1	2	3	4	1x3	1x4
28/fev	665,57358	2,16	4,6731	2,473	9,140	5,3466	19,7583
1-mar	665,02174	2,47	6,1172	9,140	-4,527	22,6060	-11,1959
2-mar	665,33330	9,14	83,5396	-4,527	-4,360	-41,3740	-39,8504
3-mar	672,00000	-4,53	20,4910	-4,360	0,640	19,7364	-2,8971
4-mar	658,33330	-4,36	19,0096	0,640	-2,860	-2,7904	12,4696
5-mar	658,50000	0,64	0,4096	-2,860	-12,027	-1,8304	-7,6971
6-mar	663,50000	-2,86	8,1796	-12,027	0,807	34,3963	-2,3071
7-mar	660,00000	-12,03	144,6407	0,807	1,723	-9,7015	-20,7260
8-mar	650,83333	0,81	0,6507	1,723	2,640	1,3902	2,1296
9-mar	663,66667	1,72	2,9699	2,640	-12,860	4,5496	-22,1621
10-mar	664,58333	2,64	6,9697	-12,860	3,140	-33,9507	8,2897
11-mar	665,50002	-12,86	165,3796	3,140	1,807	-40,3804	-23,2342
12-mar	650,00000	3,14	9,8596	1,807	-4,773	5,6730	-14,9874
13-mar	666,00000	1,81	3,2642	-4,773	1,205	-8,6235	2,1775
14-mar	664,66670	-4,77	22,7822	1,205	3,140	-5,7527	-14,9874
15-mar	658,08693	1,21	1,4526	3,140	-2,860	3,7845	-3,4470
16-mar	664,06524	3,14	9,8596	-2,860	6,640	-8,9804	20,8496
17-mar	666,00000	-2,86	8,1796	6,640	-0,527	-18,9904	1,5064
18-mar	660,00000	6,64	44,0896	-0,527	-1,193	-3,4973	-7,9235
19-mar	669,50000	-0,53	0,2774	-1,193	12,473	0,6285	-6,5697
20-mar	662,33330	-1,19	1,4240	12,473	-12,943	-14,8844	15,4453
21-mar	661,66670	12,47	155,5832	-12,943	26,140	-161,4461	326,0521
22-mar	675,33330	-12,94	167,5299	26,140	7,807	-338,3387	-101,0443
23-mar	649,91667	26,14	683,2996	7,807	-1,860	204,0663	-48,6204
24-mar	689,00000	7,81	60,9440	-1,860	3,640	-14,5204	28,4163

25-mar	670,66667	-1,86	3,4596	3,640	4,473	-6,7704	-8,3203
26-mar	661,00000	3,64	13,2496	4,473	4,640	16,2828	16,8896
27-mar	666,50000	4,47	20,0104	4,640	-4,693	20,7561	-20,9947
28-mar	667,33330	4,64	21,5296	-4,693	-0,193	-21,7771	-0,8969
29-mar	667,50000	-4,69	22,0274	-0,193	-5,193	0,9072	24,3740
30-mar	658,16667	-0,19	0,0374	-5,193	-2,860	1,0039	0,5528
31-mar	662,66670	-5,19	26,9707	-2,860	-12,860	14,8529	66,7862
1-abr	657,66667	-2,86	8,1796	-12,860	1,140	36,7796	-3,2604
2-abr	660,00000	-12,86	165,3796	1,140	0,010	-14,6604	-0,1230
3-abr	650,00000	1,14	1,2996	0,010	-1,664	0,0109	-1,8973
4-abr	664,00000	0,01	0,0001	-1,664	6,640	-0,0159	0,0635
5-abr	662,86957	-1,66	2,7700	6,640	0,473	-11,0511	-0,7877
6-abr	661,19568	6,64	44,0896	0,473	-10,193	3,1427	-67,6837
7-abr	669,50000	0,47	0,2240	-10,193	-6,860	-4,8245	-3,2468
8-abr	663,333	-10,19	103,9040	-6,860	-11,860	69,9262	120,8929
9-abr	652,667	-6,86	47,0596	-11,860	-9,360	81,3596	64,2096
10-abr	656,00000	-11,86	140,6596	-9,360	2,140	111,0096	-25,3804
11-abr	651,000	-9,36	87,6096	2,140	6,473	-20,0304	-60,5901
12-abr	653,50000	2,14	4,5796	6,473	-7,860	13,8529	-16,8204
13-abr	665,000	6,47	41,9036	-7,860	15,075	-50,8801	97,5836
14-abr	669,333	-7,86	61,7796	15,075	-3,219	-118,4878	25,2989
15-abr	655,00000	15,07	227,2491	-3,219	-0,708	-48,5211	-10,6703
16-abr	677,93478	-3,22	10,3600	-0,708	1,444	2,2783	-4,6489
17-abr	659,64130	-0,71	0,5010	1,444	7,140	-1,0223	-5,0539
18-abr	662,15217	1,44	2,0861	7,140	1,140	10,3126	1,6466
19-abr	664,30435	7,14	50,9796	1,140	-7,860	8,1396	-56,1204
20-abr	670,00000	1,14	1,2996	-7,860	-0,360	-8,9604	-0,4104
21-abr	664,00000	-7,86	61,7796	-0,360	8,618	2,8296	-67,7395
22-abr	655,00000	-0,36	0,1296	8,618	15,792	-3,1026	-5,6852
23-abr	662,50000	8,62	74,2744	15,792	5,591	136,1011	48,1866
24-abr	671,47826	15,79	249,3928	5,591	-11,360	88,2975	-179,3991
25-abr	678,65217	5,59	31,2617	-11,360	-7,403	-63,5163	-41,3943
26-abr	668,45122	-11,36	129,0496	-7,403	-3,530	84,1032	40,1008
27-abr	651,50000	-7,40	54,8111	-3,530		26,1342	
28-abr	655,45655	-3,53	12,4609				
29-abr	659,33000						
28/fev	665,57358	2,16	4,6731	2,473	9,140	5,3466	19,7583
1-mar	665,02174	2,47	6,1172	9,140	-4,527	22,6060	-11,1959
soma	40453,91		2492,6318			546,6458	-702,1898